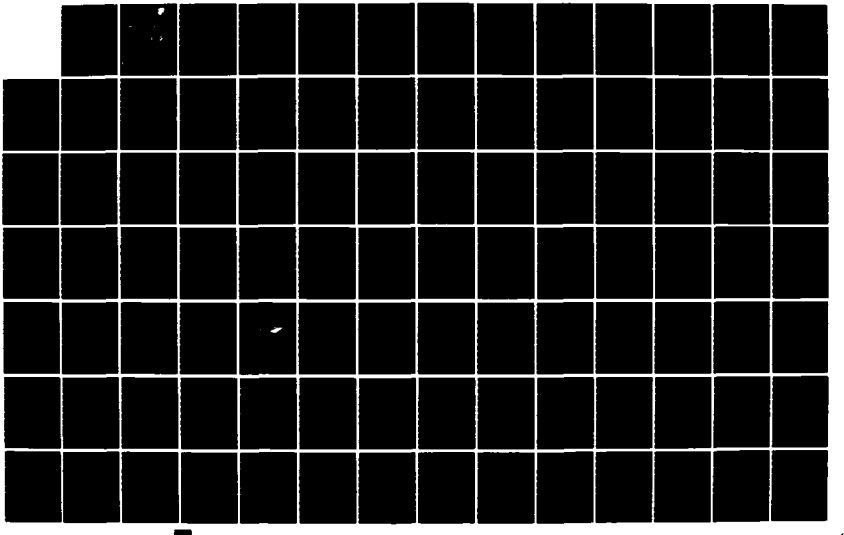
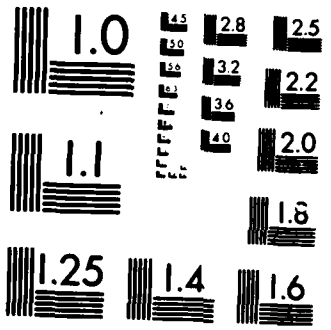
AD-A165 838 STABLE COMPENSATION OF NONLINEAR COMMUNICATIONS SYSTEMS (USING VOLTERRA S..(U) UNIVERSITY OF SOUTH FLORIDA TAMPA DEPT OF ELECTRICAL ENGINEER.. V K JAIN ET AL. 1/2

UNCLASSIFIED DEC 85 RADC-TR-85-243-VOL-2 F/G 17/2 NL





RADC-TR-85-243, Vol II (of three)
Final Technical Report
December 1985

STABLE COMPENSATION OF NONLINEAR COMMUNICATIONS SYSTEMS (USING VOLTERRA SYSTEMS CHARACTERIZATION)

University of South Florida

V. K. Jain and T. E. McClellan

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-85-243, Vol II (of three) has been reviewed and is approved for publication.

APPROVED: Daniel J. Kenneally

DANIEL J. KENNEALLY
Project Engineer

APPROVED:

W. S. TUTHILL, COLONEL, USAF
Chief, Reliability & Compatibility Division

FOR THE COMMANDER: John A. Ritz

JOHN A. RITZ
Plans & Programs Division

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (RBCT) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

ADA 165838

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS N/A
2a SECURITY CLASSIFICATION AUTHORITY N/A	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) N/A	5 MONITORING ORGANIZATION REPORT NUMBER(S) RADC-TR-85-243, Vol II (of three)
6a NAME OF PERFORMING ORGANIZATION University of South Florida 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Rome Air Development Center (RBCT)
6c ADDRESS (City, State, and ZIP Code) Department of Electrical Engineering Tampa FL 33620	7b ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700
8a NAME OF FUNDING/SPONSORING ORGANIZATION Rome Air Development Center 8b OFFICE SYMBOL (If applicable) RBCT	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F30602-82-C-0135
8c ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 62702F; PROJECT NO: 2338; TASK NO: 03; WORK UNIT ACCESSION NO: 42

11 TITLE (Include Security Classification)
STABLE COMPENSATION OF NONLINEAR COMMUNICATIONS SYSTEMS (USING VOLTERRA SYSTEMS CHARACTERIZATION)

12 PERSONAL AUTHOR(S)
V. K. Jain, T. E. McClellan

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM Jan 84 TO Apr 85	December 1985	102

16 SUPPLEMENTARY NOTATION
N/A

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	
09	03		Communication Systems, Intermodulation,
09	01		Frequency-Dependent Nonlinearity, IM Interference,
			Volterra Characterization, Compensator

19 ABSTRACT (Continue on reverse if necessary and identify by block number)
Using a mean-square IM criterion, a new approach for the reduction of intermodulation distortion is developed here. Advantages realized are: (1) reduction of IM distortion by 15 to 50 dB, (2) guaranteed stability of the compensator, (3) a minimum pre-specified damping ration of the blocks of the nonlinear compensator. A design program based on the approach is also presented.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL Daniel J. Kenneally	22b TELEPHONE (Include Area Code) (315) 330-2519
	22c OFFICE SYMBOL RADC (RBCT)

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

17. COSATI CODES (Continued)

Field	Group
17	04

18. SUBJECT TERMS (Continued)

Nonlinear Transfer Function
Associated Response Frequency Space
Post-Compensator
Computer Program
Reparametrization
Compensator

ACKNOWLEDGEMENTS

The authors wish to thank Mr. D. J. Kenneally of Rome Air Development Center for his helpful comments and criticisms of this work. His comments on the research initiatives at RADC for C^3I system interference suppression, based on Volterra identification and compensations, have lent invaluable insights to this research effort.

All experimental work in this report was done at the Communications and Signal Processing Laboratory of the University of South Florida.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

All communication channels suffer degradation from nonlinearities. Analysis and correction of the nonlinear distortion thus produced is seldom feasible with just linear methods, since they fail to capture all of the phenomena involved. Because of this, the **Nonlinear Transfer Function** (NLTF) approach, based on the Volterra theory of nonlinear systems [1], has recently been applied for weakly nonlinear systems [2]. A weakly nonlinear system is one for which the response is dominated by the linear contributions for (amplitude) ranges of the input which are encountered in practice. Such systems are also sometimes referred to as quasi-linear systems.

In communication systems many sources for nonlinear distortion exist from nonlinear discrete devices in amplifiers, metal-to-metal oxide junctions in aircraft shells, antenna structures, etc. The susceptibility to interference can be especially pronounced when the transmitter-receiver pair is located in an electronically dense command platform having a profusion of RF emitters and receptors. For example, when multiple carriers are amplified simultaneously by one transmitter, **Intermodulation** (IM) products are generated due to the nonlinearities in the power amplifier (TWT or Klystron). Similarly, a strong (locally) transmitted signal leaking into a receptor can, when processed simultaneously with a weak but desired received carrier by a nonlinear

element, produce intermodulation effects.

This report introduces a practical method for compensating the undesirable nonlinear effects produced within such a communication link. Using the known characteristics of the channel and an appropriately selected post-compensator structure, optimum parameter values are found for the compensator. The resulting design achieves significant reduction of nonlinear effects.

As stated above, we cast the compensation problem in terms of a **Mean-Square** (MS) intermodulation criterion. This MS-IM criterion, formulated for a frequency band of interest, is minimized by the design program VCOMP3. This program yields the optimum compensator parameters. High reliability in the minimization process is achieved by use of a powerful optimization package NL2SNO [4]. Clever parametrization of the block **Transfer Functions** (TF) guarantees the stability of the complete compensator over the entire parameter space. Examples provided in the report demonstrate both the simplicity and the high degree of effectiveness achievable through this new methodology.

CHAPTER II

REVIEW OF VOLTERRA-BASED SYSTEMS

In the analysis of wide band amplifiers, it is often assumed that the output depends only on the input signal applied at the same instant of time. The input/output relation can thus be expressed with a power series expansion as follows:

$$y(t) = a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + \cdots \tag{1}$$

where $x(t)$ and $y(t)$ denote the input and output signals, respectively, and the coefficients a_n are time-independent constants. In general, however, the output $y(t)$ is also dependent on the past input signal. An output expression that includes the contribution of the previous input variations is the sum of multidimensional convolution integrals

$$y(t) = \sum_{n=1}^{\infty} y_n(t) \tag{2}$$

where

$$y_n(t) = \int_{-\infty}^{\infty} \int \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \tau_2, \cdots, \tau_n) x(t-\tau_1) x(t-\tau_2) \cdots x(t-\tau_n) d\tau_1 d\tau_2 \cdots d\tau_n \tag{3}$$

and h_n is a real-valued symmetric function of n real variables.

Expression (2) is usually referred to as the Volterra series. This representation shows that a nonlinear system may be regarded as the combination of a linear and a number of higher order nonlinear

subsystems. Each of these subsystems is characterized by a n-dimensional impulse response $h_n(\tau_1,\tau_2,\cdots,\tau_n)$, also called the nth order Volterra kernel [1]. For a physically realizable system, h_n is causal, i.e., it has the value zero whenever any of its arguments is negative. Also, these kernels must be absolutely summable for a stable system. The nth order transfer function is defined as the n-fold Laplace transform of h_n, i.e.,

$$H_n(s_1,\cdots,s_n) = \int_{-\infty}^{\infty}\int\cdots\int_{-\infty}^{\infty} h_n(\tau_1,\tau_2,\cdots,\tau_n)e^{-(s_1\tau_1+\cdots+s_n\tau_n)} \cdot d\tau_1 d\tau_2\cdots d\tau_n \qquad (4)$$

In particular, we shall call $H_1(s_1)$ the linear transfer function. In Fig. 1 a block diagram is shown representing the multi-order responses summed to produce the final system response.

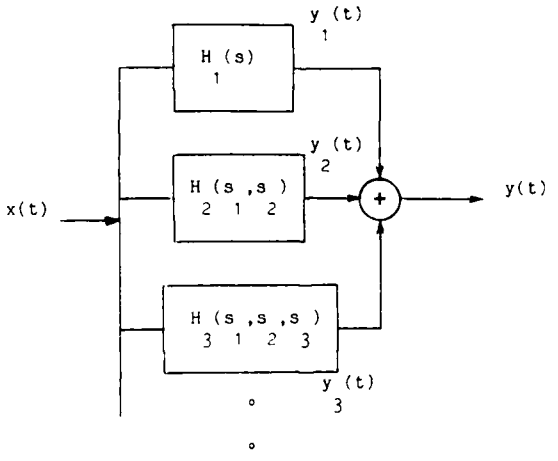


Fig. 1. Model of a weakly nonlinear channel using the Volterra series representation.

VOLTERRA REPRESENTATION OF THE QUADRATIC SYSTEM

In this section we will discuss in detail the second-order (i.e., quadratic) Volterra subsystems. Fig. 2 shows an example of a quadratic subsystem where $H_a(s)$, $H_b(s)$ and $H_c(s)$ are linear transfer functions. This particular example represents the most basic second-order subsystem and can be shown [2] to be characterized by the quadratic transfer function

$$H_2(s_1,s_2) = H_a(s_1)H_b(s_2)H_c(s_1+s_2) \tag{5}$$

Note that in general

$$\tilde{h}_2(\tau_1,\tau_2) \neq \tilde{h}_2(\tau_2,\tau_1)$$

where $\tilde{h}_2$ denotes the unsymmetric form of the kernel. The unsymmetric kernel can be symmetrized by defining a new kernel as

$$h_2(\tau_1,\tau_2) = \frac{1}{2}\left[\tilde{h}_2(\tau_1,\tau_2) + \tilde{h}_2(\tau_2,\tau_1)\right]$$

Since it is not possible to measure $\tilde{h}_2(\tau_1,\tau_2)$ from only the system input and output [1], it makes sense to consider only symmetric kernels in our analysis without any loss of generality. In particular, we shall assume that $H_a(s) = H_b(s)$ for the second-order subsystem shown in Fig. 2. Correspondingly the time-domain Volterra kernel can be shown to be [2]

$$h_2(\sigma_1,\sigma_2) = \int_{-\infty}^{\infty} h_c(\tau)h_a(\sigma_1-\tau)h_b(\sigma_2-\tau)d\tau \tag{6}$$

Associated Two-dimensional Response

The expression

$$y_{(2)}(t_1,t_2) = \int_{-\infty}^{\infty}\int h_2(\tau_1,\tau_2)x(t_1-\tau_1)x(t_2-\tau_2)d\tau_1 d\tau_2 \tag{7}$$

is known as the associated two-dimensional response. Its significance

arises from the fact that if we set $t_1=t_2=t$ then,

$$y_{(2)}(t,t) = y_2(t) \tag{8}$$

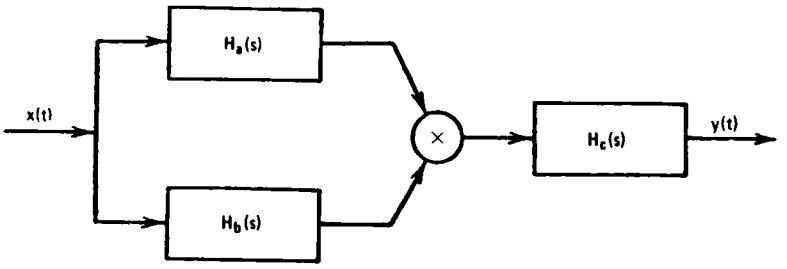


Fig. 2. Basic second-order Volterra system.

That is, $y_{(2)}(t_1,t_2)$ contains the response $y_2(t)$ along the 45^o line in the t_1-t_2 plane. Further, this associated response can be computed through linear operations on the input signal. Indeed, the two-dimensional Fourier transform of $y_{(2)}(t_1,t_2)$ is

$$Y_{(2)}(\omega_1,\omega_2) = H_2(\omega_1,\omega_2)X(\omega_1)X(\omega_2) \tag{9}$$

where $X(\omega)$ is the Fourier transform of the input signal. Taking the two-dimensional inverse Fourier transform yields

$$y_{(2)}(t_1,t_2) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int H_2(\psi_1,\psi_2)X(\psi_1)X(\psi_2) \cdot e^{j(\psi_1 t_1+\psi_2 t_2)}\,d\psi_1\,d\psi_2 \tag{10}$$

Hence, by setting $t_1=t_2=t$, we have

$$y_2(t) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty}\int H_2(\psi_1,\psi_2)X(\psi_1)X(\psi_2)e^{j(\psi_1+\psi_2)t}\,d\psi_1 d\psi_2 \qquad (11)$$

Further, by taking the Fourier transform of (11) we obtain

$$Y_2(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty}\int H_2(\psi_1,\psi_2)X(\psi_1)X(\psi_2)\delta(\omega-\psi_1-\psi_2)d\psi_1 d\psi_2 \qquad (12)$$

In view of (9) we can also write

$$Y_2(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty}\int Y_{(2)}(\psi_1,\psi_2)\delta(\omega-\psi_1-\psi_2)d\psi_1 d\psi_2$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_{(2)}(\psi_1,\omega-\psi_1)d\psi_1 \qquad (13)$$

In other words, $Y_2(\omega)$ is the integral of $Y_{(2)}(\psi_1,\psi_2)$ along the line $\omega=\psi_1+\psi_2$; alternatively, $Y_2(\omega)$ is the profile of $Y_{(2)}(\psi_1,\psi_2)$ along the line $\psi_1=\psi_2$.

Let us now consider two simple examples in order to better appreciate the intermediate use of the associated response for finding the nonlinear system response.

Example 1

Let us compute the quadratic system response to a single complex sinusoid,

$$x(t) = Ae^{j\Omega_1 t} \qquad (14)$$

Translated into the frequency domain the input is written as

$$X(\omega) = 2\pi A\delta(\omega-\Omega_1) \qquad (15)$$

Then,

$$Y_{(2)}(\omega_1,\omega_2) = (2\pi)^2 H_2(\omega_1,\omega_2)\delta(\omega_1-\Omega_1)\delta(\omega_2-\Omega_1)$$

$$Y_2(\omega) = A^2 \int_{-\infty}^{\infty}\int H_2(\psi_1,\psi_2)\delta(\psi_1-\Omega_1)\delta(\psi_2-\Omega_1)\delta(\omega-\psi_2-\psi_1)d\psi_2 d\psi_1$$

$$= \frac{A^2}{2\pi} \int_{-\infty}^{\infty} H_2(\psi_1,\Omega_1)\delta(\psi_1-\Omega_1)\delta(\omega-\psi_1-\Omega_1)d\psi_1$$

$$= A^2 H_2(\Omega_1,\Omega_1)\delta(\omega-2\Omega_1)$$

$$y_2(t) = A^2 H_2(\Omega_1,\Omega_1)e^{j2\Omega_1 t} \tag{16}$$

We note that the output frequency of the quadratic nonlinearity is twice the input sinusoids's frequency.

Example 2

Suppose now that the input to the quadratic system consists of the sum of two sinusoids, given by

$$x(t) = A_1 e^{j\Omega_1 t} + A_2 e^{j\Omega_2 t} \tag{17}$$

$$X(\omega) = 2\pi[A_1\delta(\omega-\Omega_1)+A_2\delta(\omega-\Omega_2)]$$

The corresponding output is

$$Y_2(\omega) = \int_{-\infty}^{\infty} H_2(\psi_1,\psi_2)[A_1\delta(\psi_1-\Omega_1)+A_2\delta(\psi_1-\Omega_2)]$$

$$\cdot [A_1\delta(\psi_2-\Omega_1)+A_2\delta(\psi_2-\Omega_2)]\delta(\omega-\psi_1-\psi_2)d\psi_1 d\psi_2$$

$$y_2(t) = A_1^2 H_2(\Omega_1,\Omega_1)e^{j2\Omega_1 t} + 2A_1A_2H_2(\Omega_1,\Omega_2)e^{j(\Omega_1+\Omega_2)t}$$

$$+A_2^2 H_2(\Omega_2,\Omega_2)e^{j2\Omega_2 t} \tag{18}$$

Thus an input consisting of the sum of two sinusoids results in an output containing three frequencies.

The results of the above examples can be generalized to the case of an input consisting of the sum of M sinusoids, expressed by

$$x(t) = \sum_{i=1}^{M} A_i e^{j\Omega_i t} \tag{19}$$

The quadratic response to this input can be shown to be

$$y_2(t) = \sum_{1 \le k \le i}^{M} \sum^{M} \frac{2!\, A_i A_k}{m_1! m_2! \cdots m_M!} H_2(\Omega_i, \Omega_k) e^{j(\Omega_i + \Omega_k)t} \qquad (20)$$

where

$$m_1 + m_2 + \cdots + m_M = 2$$

and m_ℓ is the number of occurrences of the ℓth frequency (as arguments of H_2).

SECOND-ORDER VOLTERRA SYSTEM RESPONSE TO THE SUM OF M REAL SINUSOIDS

Multi-channel communication systems must necessarily employ multiple carrier waveforms. In a linear environment no new frequencies are produced which were not present in the input. Nonlinearities, even of a mild nature, will generate new terms. We derive a closed form expression for the response of a second-order system to an input consisting of the sum of M sinusoids. Given the second-order symmetric nonlinear system of Fig. 3, let the input be the sum of M cosine waveforms

$$x(t) = A_1 \cos(\Omega_1 t) + A_2 \cos(\Omega_2 t) + \cdots + A_M \cos(\Omega_M t). \qquad (21)$$

The output at v_a is denoted by

$$\begin{aligned} v_a(t) &= v_{a_1} + v_{a_2} + \cdots + v_{a_M} \\ &= A_1 |H_a(\Omega_1)| \cos(\Omega_1 t + \theta_{\Omega_1}) + A_2 |H_a(\Omega_2)| \cos(\Omega_2 t + \theta_{\Omega_2}) + \cdots \\ &\quad + A_M |H_a(\Omega_M)| \cos(\Omega_M t + \theta_{\Omega_M}) . \end{aligned} \qquad (22)$$

where θ_{Ω_i} is defined as Phase$\{H_a(\Omega_i)\}$

To proceed further with the analysis, let us compute the response at w,

$$w(t) = v_a \cdot v_a$$

$$= \sum_{i=1}^{M} \sum_{k=1}^{M} v_{a_i} v_{a_k}$$

$$= \sum_{1 \le i \le k}^{M} \sum^{M} \frac{2!}{m_1! m_2! \cdots m_M!} V_{a_i} V_{a_k}$$

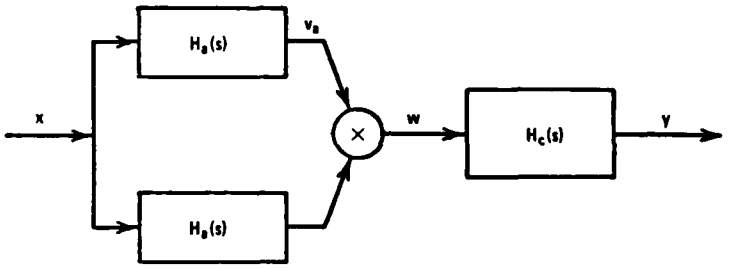


Fig. 3. Basic second-order nonlinear system.

$$= \sum_{1 \le i \le k}^{M} \sum^{M} \Gamma_{ik} |H_a(\Omega_i)||H_a(\Omega_k)| \cos(\Omega_i t + \theta_{\Omega_i}) \cos(\Omega_k t + \theta_{\Omega_k}) \qquad (23)$$

where

$$\Gamma_{ik} \triangleq \frac{2! A_i A_k}{m_1! m_2! \cdots m_M!}$$

and $\theta_{\Omega_i} \triangleq \text{Phase}\{H_a(\Omega_i)\}$

Equation (23) may be simplified by using the trigonometric identity

$$\cos(\xi_1 t)\cos(\xi_2 t) = \frac{1}{2}\left[\cos(\xi_1+\xi_2)t + \cos(\xi_1-\xi_2)t\right]$$

Thus, we have

$$w(t) = \frac{1}{2} \sum_{1 \le k \le M} \sum_{k \le i \le M} \Gamma_{ik} |H_a(\Omega_i) H_a(\Omega_k)| \left[\cos((\Omega_i+\Omega_k)t + \theta_{\Omega_i} + \theta_{\Omega_k}) + \cos((\Omega_i-\Omega_k)t + \theta_{\Omega_i} - \theta_{\Omega_k}) \right] \quad (24)$$

The output is then written as

$$y(t) = \frac{1}{2} \sum_{1 \le k \le M} \sum_{k \le i \le M} \Gamma_{ik} |H_a(\Omega_i) H_b(\Omega_k)| \cdot \left[|H_c(\Omega_i+\Omega_k)| \cos((\Omega_i+\Omega_k)t + \theta_{\Omega_i} + \theta_{\Omega_k} + \phi_{\Omega_i+\Omega_k}) + |H_c(\Omega_i-\Omega_k)| \cos((\Omega_i-\Omega_k)t + \theta_{\Omega_i} - \theta_{\Omega_k} + \phi_{\Omega_k-\Omega_i}) \right] \quad (25)$$

To illustrate the use of this consider an input consisting of the sum of two sinusoids,

$$x(t) = A_1 \cos(\Omega_1 t) + A_2 \cos(\Omega_2 t)$$

The output is then

$$y(t) = \frac{1}{2} A_1^2 |H_a(\Omega_1)|^2 |H_c(2\Omega_1)| \left[\cos(2\Omega_1 t + \theta_{2\Omega_1} + 2\phi_{\Omega_1}) + 1 \right] + A_1 A_2 |H_a(\Omega_1) H_a(\Omega_2) H_c(\Omega_1+\Omega_2)| \left[\cos((\Omega_1+\Omega_2)t + \theta_{\Omega_1} + \theta_{\Omega_2} + \phi_{\Omega_1+\Omega_2}) + \cos((\Omega_1-\Omega_2)t + \theta_{\Omega_1} - \theta_{\Omega_2} + \phi_{\Omega_1-\Omega_2}) \right] + \frac{1}{2} A_2^2 |H_a(\Omega_2)|^2 |H_c(2\Omega_2)| \left[\cos(2\Omega_2 t + 2\theta_{\Omega_2} + \phi_{2\Omega_2}) + 1 \right] \quad (26)$$

where $\phi_{\Omega_i} \triangleq \text{Phase}\{H_c(\Omega_i)\}$

Since $H_2(\Omega_1, \Omega_2) = H_a(\Omega_1) H_a(\Omega_2) H_c(\Omega_1+\Omega_2)$, the above expression simplifies to

$$y(t) = \frac{1}{2} A_1^2 \left[|H_2(\Omega_1, \Omega_1)| \cos(2\Omega_1 t + \alpha_1) + H_2(\Omega_1, -\Omega_1) \right]$$

$$+A_1A_2\left[|H_2(\Omega_1,\Omega_2)|\cos((\Omega_1+\Omega_2)t+\alpha_2)\right.$$

$$\left.+|H_2(\Omega_1,-\Omega_2)|\cos((\Omega_1-\Omega_2)t+\alpha_3)\right]$$

$$+\frac{1}{2}A_2^2\left[|H_2(\Omega_2,\Omega_2)|\cos(2\Omega_2t+\alpha_4)+|H_2(\Omega_2,-\Omega_2)|\right] \quad (27)$$

where

$$\alpha_1 \triangleq \text{Phase}\{H_2(\Omega_1,\Omega_1)\}$$

$$\alpha_2 \triangleq \text{Phase}\{H_2(\Omega_1,\Omega_2)\}$$

$$\alpha_3 \triangleq \text{Phase}\{H_2(\Omega_1,-\Omega_2)\}$$

$$\alpha_4 \triangleq \text{Phase}\{H_2(\Omega_2,\Omega_2)\}$$

THIRD-ORDER NONLINEAR RESPONSE TO A SUM OF MULTIPLE COSINE CARRIERS

The cubic nonlinear response to a sum of M cosine carrier waveforms is computed as above. Given the nonlinear system of Fig. 4 the response to an input of M cosines is

$$y(t) = \frac{1}{4}\sum_{i}^{M}\sum_{k\le}^{M}\sum_{\ell\le}^{M}\Gamma_{ik\ell}\,|Q_{ik\ell}|$$

$$\cdot\left[|R_{i,-k,-\ell}|\cos((\Omega_i-\Omega_k-\Omega_\ell)t+\theta_{\Omega_i}-\theta_{\Omega_k}-\theta_{\Omega_\ell}+\phi_{\Omega_i,-\Omega_k,-\Omega_\ell})\right.$$

$$+|R_{i,k,-\ell}|\cos((\Omega_i-\Omega_k-\Omega_\ell)t+\theta_{\Omega_i}+\theta_{\Omega_k}-\theta_{\Omega_\ell}+\phi_{\Omega_i,\Omega_k,-\Omega_\ell})$$

$$+|R_{i,-k,\ell}|\cos((\Omega_i-\Omega_k+\Omega_\ell)t+\theta_{\Omega_i}-\theta_{\Omega_k}+\theta_{\Omega_\ell}+\phi_{\Omega_i,-\Omega_k,\Omega_\ell})$$

$$\left.+|R_{i,k,1}|\cos((\Omega_i+\Omega_k+\Omega_\ell)t+\theta_{\Omega_i}+\theta_{\Omega_k}+\theta_{\Omega_\ell}+\phi_{\Omega_i,\Omega_k,\Omega_\ell})\right] \quad (28)$$

where $\Gamma_{ik\ell} \triangleq \dfrac{3!A_iA_kA_\ell}{m_1!m_2!\cdots m_M!}$

$$Q_{ik\ell} \triangleq H_a(\Omega_i)H_a(\Omega_k)H_a(\Omega_\ell)$$

$$R_{ik\ell} \triangleq H_d(\Omega_i+\Omega_k+\Omega_\ell)$$

$$\theta_{\Omega_i} \triangleq \text{Phase}\{H_a(\Omega_i)\}$$

$$\phi_{\Omega_i,\Omega_k,\Omega_\ell} \triangleq \text{Phase}\{H_d(\Omega_i+\Omega_k+\Omega_\ell)\}$$

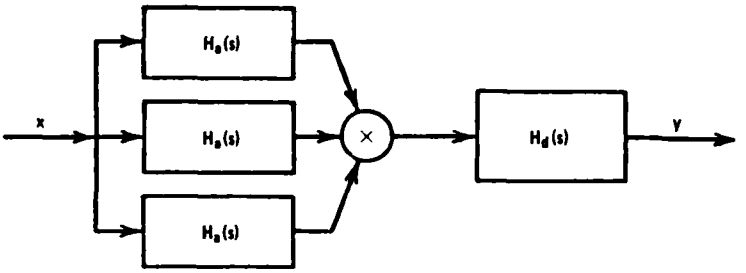


Fig. 4. Third-order nonlinear system.

This expression for the third-order response can be expressed more concisely with the following notation,

$$y(t) = \frac{1}{4} \sum_{i}^{M} \sum_{k \le}^{M} \sum_{\ell}^{M} \Gamma_{ik\ell}\, Q_{ik\ell} |H_d(\Omega_i \pm \Omega_k \pm \Omega_\ell)| \cdot \cos((\Omega_i \pm \Omega_k \pm \Omega_\ell)t + \theta_{\Omega_i} \pm \theta_{\Omega_k} \pm \theta_{\Omega_\ell} + \phi_{\Omega_i \pm \Omega_k \pm \Omega_\ell}) \quad (29)$$

To give the reader a feel for the complexity of the cubic nonlinearity response to a sum of sinusoids, consider an input consisting of the sum of three cosine waveforms, with distinct frequencies. In the real-world such a signal might arise from an amplitude modulated carrier, wherein the modulation signal is a low frequency sinusoid; and H_3 might be the

third-order nonlinear transfer function of the power amplifier. The input is

$$x(t) = A_1\cos(\Omega_1 t)+A_2\cos(\Omega_2 t)+A_3\cos(\Omega_3 t)$$

Using expression (29) the output is

$$\begin{aligned}
y(t) = \tfrac{1}{4}\{&\Gamma_{111}|Q_{111}|[3|H_a(\Omega_1)|\cos((\Omega_1)t+\theta_{\Omega_1}) \\
&+|H_d(3\Omega_1)|\cos((3\Omega_1)t+3\theta_{\Omega_1}+\phi_{3\Omega_1})] \\
&+\Gamma_{112}|Q_{112}|[2|H_a(\Omega_2)|\cos((\Omega_2)t+2\theta_{\Omega_2}) \\
&+|H_d(2\Omega_1-\Omega_2)|\cos((2\Omega_1-\Omega_2)t+2\theta_{\Omega_1}-\theta_{\Omega_2}+\phi_{2\Omega_1-\Omega_2}) \\
&+|H_d(2\Omega_1+\Omega_2)|\cos((2\Omega_1+\Omega_2)t+2\theta_{\Omega_1}+\theta_{\Omega_2}+\phi_{2\Omega_1+\Omega_2})] \\
&+\Gamma_{122}|Q_{122}||[|H_d(\Omega_1-2\Omega_2)|\cos((\Omega_1-2\Omega_2)t+\theta_{\Omega_1}-2\theta_{\Omega_2}+\phi_{\Omega_1-2\Omega_2}) \\
&+2|H_a(\Omega_1)|\cos((\Omega_1)t+2\theta_{\Omega_1}) \\
&+|H_d(\Omega_1+2\Omega_2)|\cos((\Omega_1+2\Omega_2)t+\theta_{\Omega_1}+2\theta_{\Omega_2}+\phi_{\Omega_1+2\Omega_2})] \\
&+\Gamma_{222}|Q_{222}|[3|H_a(\Omega_2)|\cos((\Omega_2)t+\theta_{\Omega_2}) \\
&+|H_d(3\Omega_2)|\cos((3\Omega_2)t+3\theta_{\Omega_2}+\phi_{3\Omega_2}] \\
&+\Gamma_{113}|Q_{113}|[2|H_a(\Omega_3)|\cos((\Omega_3)t+2\theta_{\Omega_3}) \\
&+|H_d(2\Omega_1+\Omega_3)|\cos((2\Omega_1+\Omega_3)t+2\theta_{\Omega_1}+\theta_{\Omega_3}+\phi_{2\Omega_1+\Omega_3}) \\
&+|H_d(2\Omega_1-\Omega_3)|\cos((2\Omega_1-\Omega_3)t+\theta_{2\Omega_1}-\theta_{\Omega_3}+\phi_{2\Omega_1-\Omega_3})] \\
&+\Gamma_{123}|Q_{123}|[|H_d(\Omega_1-\Omega_2-\Omega_3)|\cos((\Omega_1-\Omega_2-\Omega_3)t+\theta_{\Omega_1}-\theta_{\Omega_2}-\theta_{\Omega_3}+\phi_{\Omega_1-\Omega_2-\Omega_3}) \\
&+|H_d(\Omega_1+\Omega_2-\Omega_3)|\cos((\Omega_1+\Omega_2-\Omega_3)t+\theta_{\Omega_1}+\theta_{\Omega_2}-\theta_{\Omega_3}+\phi_{\Omega_1+\Omega_2-\Omega_3}) \\
&+|H_d(\Omega_1-\Omega_2+\Omega_3)|\cos((\Omega_1-\Omega_2+\Omega_3)t+\theta_{\Omega_1}-\theta_{\Omega_2}+\theta_{\Omega_3}+\phi_{\Omega_1-\Omega_2+\Omega_3}) \\
&+|H_d(\Omega_1+\Omega_2+\Omega_3)|\cos((\Omega_1+\Omega_2+\Omega_3)t+\theta_{\Omega_1}+\theta_{\Omega_2}+\theta_{\Omega_3}+\phi_{\Omega_1+\Omega_2+\Omega_3})] \\
&+\Gamma_{133}|Q_{133}|[H_d(\Omega_1-2\Omega_3)|\cos((\Omega_1-2\Omega_3)t+\theta_{\Omega_1}-2\theta_{\Omega_3}+\phi_{\Omega_1-2\Omega_3})
\end{aligned}$$

$$+|H_d(\Omega_1+\Omega_2+\Omega_3)|\cos((\Omega_1+\Omega_2+\Omega_3)t+\theta_{\Omega_1}+\theta_{\Omega_2}+\theta_{\Omega_3}+\phi_{\Omega_1+\Omega_2+\Omega_3})]$$

$$+\Gamma_{133}|Q_{133}|[H_d(\Omega_1-2\Omega_3)|\cos((\Omega_1-2\Omega_3)t+\theta_{\Omega_1}-2\theta_{\Omega_3}+\phi_{\Omega_1-2\Omega_3})$$

$$+2|H_a(\Omega_1)|\cos((\Omega_1)t+2\theta_{\Omega_1})$$

$$+|H_d(\Omega_1+2\Omega_3)|\cos((\Omega_1+2\Omega_3)t+\theta_{\Omega_1}+2\theta_{\Omega_3}+\phi_{\Omega_1+2\Omega_3})]$$

$$+\Gamma_{223}|Q_{223}|[2|H_a(\Omega_3)|\cos((\Omega_3)t+2\theta_{\Omega_3})$$

$$+|H_d(2\Omega_2-\Omega_3)|\cos((2\Omega_2-\Omega_3)t+2\theta_{\Omega_2}-\theta_{\Omega_3}+\phi_{2\Omega_2-\Omega_3})$$

$$+|H_d(2\Omega_2+\Omega_3)|\cos((2\Omega_2+\Omega_3)t+2\theta_{\Omega_2}+\theta_{\Omega_3}+\phi_{2\Omega_2+\Omega_3})]$$

$$+\Gamma_{233}|Q_{233}|[|H_d(\Omega_2+2\Omega_3)|\cos((\Omega_2+2\Omega_3)t+\theta_{\Omega_2}+2\theta_{\Omega_3}+\phi_{\Omega_2+2\Omega_3})$$

$$+2|H_a(\Omega_2)|\cos((\Omega_2)t+2\theta_{\Omega_2})$$

$$+|H_d(\Omega_2-2\Omega_3)|\cos((\Omega_2-2\Omega_3)t+\theta_{\Omega_2}-2\theta_{\Omega_3}+\phi_{\Omega_2-2\Omega_3})]$$

$$+\Gamma_{333}|Q_{333}|[3|H_a(\Omega_3)|\cos((\Omega_3)t+\theta_{\Omega_3})$$

$$+|H_d(3\Omega_3)|\cos((3\Omega_3)t+3\theta_{\Omega_3}+\phi_{3\Omega_3})]] \qquad (30)$$

This output expression is only for the sum of three cosine inputs; if the input were the sum of seven or eight cosine waveforms, one can see the difficulty in calculating the response in terms of the frequencies generated.

Conversion of The Third-Order Volterra System Response Into The Third-Order Nonlinear Response

The response of a third-order Volterra operator can be expressed in the functional form

$$y_3(t) = H_3[x(t)]$$

in which $h_3(\tau_1,\tau_2,\tau_3)$ is the symmetric third-order Volterra kernel of the third-order Volterra operator H_3. In this section, our objective is to express the Fourier transform of the response $Y_3(\omega)$ in terms of the Fourier transform of the input $X(\omega)$ and the third-order NLTF $H_3(\omega_1,\omega_2,\omega_3)$. Note that the arguments are taken to be ω, rather than $j\omega$ to simplify the notation. Earlier, we saw how the transformation of the second-order Volterra associated response into the actual response was accomplished. We follow the same procedure for the case of the third-order Volterra operator. The third-order associated response is defined as

$$y_{(3)}(t_1,t_2,t_3) = \int_{-\infty}^{\infty}\int\int h_3(\tau_1,\tau_2,\tau_3)x(t_1-\tau_1)x(t_2-\tau_2)x(t_3-\tau_3) \cdot d\tau_1 d\tau_2 d\tau_3 \qquad (32)$$

The third-order response $y_3(t)$ is readily obtained by setting each of the arguments of $y_{(3)}$ equal to t. That is

$$y_3(t) = y_{(3)}(t,t,t) \qquad (33)$$

Let us next compute the three dimensional Fourier transform, $Y_{(3)}(\omega_1,\omega_2,\omega_3)$ of the associated function $y_{(3)}(t_1,t_2,t_3)$. Then compute $Y_3(\omega)$ from this Fourier transform of the associated response.

The three dimensional Fourier transform of the associated response is,

$$Y_{(3)}(\omega_1,\omega_2,\omega_3) = H_3(\omega_1,\omega_2,\omega_3)X(\omega_1)X(\omega_2)X(\omega_3) \qquad (34)$$

where $X(\omega)$ is the Fourier transform of the input signal. Taking the three-dimensional inverse transform yields

$$y_{(3)}(t_1,t_2,t_3) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} Y_{(3)}(\omega_1,\omega_2,\omega_3)e^{j(\omega_1 t_1+\omega_2 t_2+\omega_3 t_3)} \cdot d\omega_1 d\omega_2 d\omega_3 \quad (35)$$

Setting $t_1=t_2=t_3=t$ in our previous equation we obtain,

$$y_3(t) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} Y_{(3)}(\omega_1,\omega_2,\omega_3)e^{j(\omega_1+\omega_2+\omega_3)t} d\omega_1 d\omega_2 d\omega_3 \quad (36)$$

The inverse Fourier transform of $Y_3(\omega)$ is by definition

$$y_3(t) = \frac{1}{(2\pi)} \int_{-\infty}^{\infty} Y_3(\omega)e^{j\omega t} d\omega \quad (37)$$

Thus we wish to make a change of variable in the Fourier transform of the associated response that will give us the Fourier transform of the actual response. Let $\omega_1+\omega_2+\omega_3 = \omega$, this enables us to obtain $Y_3(\omega)$ from $Y_{(3)}(\omega_1,\omega_2,\omega_3)$. We shall make the desired change of variable in two steps. First, in eq.(38) we let $\omega_2+\omega_3=\psi_1$ to obtain

$$y_3(t) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} Y_{(3)}(\omega_1,\psi_1-\omega_3,\omega_3)e^{j(\omega_1+\psi_1)t} d\omega_1 d\psi_1 d\omega_3 \quad (38)$$

For the second change of variable, let $\psi_1+\omega_1=\omega$ in (40) to have

$$y_3(t) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} Y_{(3)}(\omega-\psi_1,\psi_1-\omega_3,\omega_3)e^{j\omega t} d\psi_1 d\omega_3 d\omega \quad (39)$$

Note that this triple integral is in the form of an inverse Fourier transform of dimension one. From this observation we can rewrite it as

$$Y_3(\omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty}\int Y_{(3)}(\omega-\psi_1,\psi_1-\omega_3,\omega_3)d\psi_1 d\omega_3 \quad (40)$$

$$= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty}\int Y_{(3)}(\omega-\psi_1,\psi_1-\psi_2,\psi_2)d\psi_1 d\psi_2 \quad (41)$$

In other words, $Y_3(\omega)$ is the integral of $Y_{(3)}(\psi_1,\psi_2,\psi_3)$ along the plane $\omega=\psi_1+\psi_2+\psi_3$; alternatively $Y_3(\omega)$ is the profile of $Y_{(3)}(\psi_1,\psi_2,\psi_3)$ along the line $\psi_1=\psi_2=\psi_3$.

CASCADING OF NONLINEAR SYSTEMS

Before we deal with the matter of compensation of nonlinear systems, we must first characterize a tandem connection of two general nonlinear systems. Fig. 5 shows the cascade connection of such Volterra systems. In this section we wish to derive an equivalent Volterra NLTFs description of the two nonlinear systems in tandem. Suppose that the two subsystems are characterized by the NLTFs H_n and P_n, $n=1,2,\ldots,\infty$.

Then

$$w = H[x] = \sum_{n=1}^{\infty} H_n[x] = \sum_{n=1}^{\infty} w_n \tag{42}$$

and the output y is given by

$$y = P[w] = \sum_{k=1}^{\infty} P_k[w] \tag{43}$$

$$= \sum_{k=1}^{\infty} P_k\Big[\sum_{n=1}^{\infty} H_n[x]\Big]$$

To use the nonlinear current method [3], let us replace x with εx to form a new output;

$$y_\varepsilon = \sum_{k=1}^{\infty} P_k\Big[\sum_{n=1}^{\infty} H_n[\varepsilon x]\Big]. \tag{44}$$

$$= \sum_{k=1}^{\infty} P_k\Big[\sum_{n=1}^{\infty} \varepsilon^n H_n[x]\Big]$$

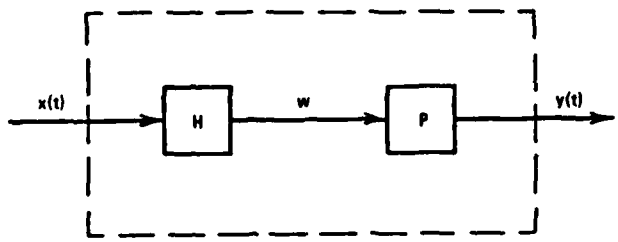


Fig. 5. Cascade of two nonlinear systems H and P.

Now, from the basic k-linear property [1] of the operator P_k, it is seen that

$$P_k\Big[\sum_{n=1}^{\infty}\varepsilon^n w_n\Big] = \sum_{n_1=1}^{\infty}\cdots\sum_{n_k=1}^{\infty}\varepsilon^{n_1}\cdots\varepsilon^{n_k}\,P_k\{w_{n_1},\cdots,w_{n_k}\} \tag{45}$$

in which $P_k\{\cdot\}$ is the k-linear Volterra operator

$$P_k\{w_{n_1},\cdots,w_{n_k}\}$$

$$= \int_{-\infty}^{\infty}\int\int\cdots\int_{-\infty}^{\infty} p_k(t_1,\cdots,t_k)w_{n_1}(t-\tau_1)\cdots w_{n_k}(t-\tau_1)d\tau_1\cdots d\tau_n \tag{46}$$

Then

$$y = \sum_{m=1}^{m}\sum_{n_1=1}^{\infty}\cdots\sum_{n_m=1}^{\infty}\varepsilon^{n_1+\cdots+n_m}\,P_m\{w_{n_1},\cdots,w_{n_m}\} \tag{47}$$

This relationship can be re-expressed in terms of the "x-y transfer function" Q as

$$y_\varepsilon = \sum_{k=1}^{\infty} \varepsilon^k \, Q_k[x] \tag{48}$$

Now if like powers of ε are collected, the Q's can be determined using "operator notation" as

$$Q_1 = P_1 H_1 \tag{49}$$

$$Q_2 = P_1 H_2 + P_2 H_1$$

$$Q_3 = P_1 H_3 + 2P_2\{H_1, H_2\} + P_3 H_1$$

$$Q_4 = P_1 H_4 + 2P_2\{H_1, H_3\} + P_2[H_2] + 3P_3\{H_1, H_1, H_2\} + P_4[H_1]$$

$$Q_5 = P_1 H_5 + 2P_2\{H_1, H_4\} + 2P_2\{H_2, H_3\} + 3P_3\{H_1, H_2, H_2\}$$
$$+ 4P_4\{H_1, H_1, H_1, H_2\} + P_5 H_1$$

$$\cdot$$

$$\cdot$$

$$Q_m = P_1 H_m + \text{terms involving } P_j \text{ for } j \leq (m-1) \tag{50}$$

This relationship is needed in our discussion of the mth-order nonlinear inverse.

These equations can be expressed in the Laplace-domain as

$$Q_1(s) = P_1(s) H_1(s)$$

$$Q_2(s_1, s_2) = P_1(s_1 + s_2) H_2(s_1, s_2) + H_1(s_1) H_1(s_2) P_2(s_1, s_2)$$

$$Q_3(s_1, s_2, s_3) =$$

$$P_1(s_1 + s_2 + s_3) H_3(s_1, s_2, s_3) + \frac{1}{3} [H_1(s_1) H_2(s_2, s_3) P_2(s_1, s_2 + s_3)$$

$$+ H_1(s_2) H_2(s_1, s_3) P_2(s_2, s_1 + s_3) + H_1(s_3) H_2(s_1, s_2) P_2(s_3, s_1 + s_2)]$$

$$+ H_1(s_1) H_1(s_2) H_1(s_3) P_3(s_1, s_2, s_3)$$

$$\cdot$$

$$\cdot \tag{51}$$

Our objective is to reduce the nonlinear response in a system without

distorting the linear characteristics. Since the linear response must not be changed, the form of Q_1 should be H_1; and therefore,

$$P_1 = I \tag{52}$$

where I is the identity operator. Furthermore, the second and third-order net nonlinear transfer functions (Q_2 and Q_3)can be made zero by choosing P_2 and P_3 as

$$P_2 = -P_1 H_2 H_1^{-1} \tag{53}$$

$$P_3 = -(H_3 + 2P_2\{H_1, H_2\})H_1^{-1}$$

Example

A simple example is used to show how a nonlinear compensator is constructed. Only the second-order distortion is reduced in this example, but in Chapter 4 we will address both second and third-order nonlinearly generated distortion. Suppose the NLTFs of a circuit with a single nonlinear device [3] are:

$$H_1(s) = Cs \frac{Ls}{LCs^2 + Lg_1 s + 1} \triangleq Cs\, Z(s) \tag{54}$$

$$H_2(s_1, s_2) = -a_2 Z(s_1+s_2)\ \{H_1(s_1)H_1(s_2)\}$$

Then P_2 becomes

$$P_2(s_1, s_2) = -H_2(s_1, s_2)H_1(s_1+s_2)^{-1}$$

$$= a_2 \frac{LC(s_1+s_2)}{LC(s_1+s_2)^2 + Lg_1(s_1+s_2) + 1} \tag{55}$$

The algebraic expression for $P_3(s_1, s_2, s_3)$ is determined by evaluating the second expression of equation (55). While this evaluation is mathematically straight forward, the resulting expression is involved.

Therefore, the specific form of P_3 will not be determined. This then completes this example.

THE APPLICATION OF VOLTERRA THEORY TO REDUCE NONLINEAR EFFECTS

Volterra theory may be used to predict and diminish undesirable effects of the nonlinearity of a system. One possible approach suggested by Wiener [1] is to use linear feedback to reduce the nonlinear contributions. Let us review the feedback method briefly.

The reduction of nonlinear response by linear feedback was applied to transistor amplifiers by Narayanan [5], In that paper the open-loop transfer function of a transistor amplifier is characterized by a sum of Volterra kernels, shown in Fig. 6. These kernels take into account the frequency-dependent nature of the nonlinearities.

The basic closed loop equations for the circuit are:

$$\begin{aligned} e &= x-v \\ y &= \mathbf{H}[e] \\ v &= \mathbf{B}[y] \\ y &= \mathbf{G}[x] \end{aligned} \tag{56}$$

The closed loop system operator **G**, expressed so that it satisfies all input signals,

$$\mathbf{G} = \mathbf{H}\,[\mathbf{I}-\mathbf{BG}] \tag{57}$$

where **I** is the identity operator. The closed loop kernels for the first, $G_1(\omega)$, second $G_2(\omega_1,\omega_2)$, and third-order, $G_3(\omega_1,\omega_2,\omega_3)$ are obtained by successively equating the linear, second and third-order terms.

$$G_1(\omega) = \frac{H_1(\omega)}{1 + B(\omega)H_1(\omega)} \tag{58}$$

$$G_2(\omega_1,\omega_2) = \frac{H_2(\omega_1,\omega_2)}{[1+H_1(\omega_1)B(\omega_1)][1+H_1(\omega_2)B(\omega_2)]} \cdot \frac{1}{1+H_1(\omega_1+\omega_2)B(\omega_1+\omega_2)} \quad (59)$$

$$\begin{aligned}\tilde{G}_3(\omega_1,\omega_2,\omega_3) = [&H_3(\omega_1,\omega_2,\omega_3) \prod_{i=1}^{3} \frac{1}{1+H_1(\omega_i)B(\omega_i)} \\ &- \{3H_2(\omega_1,\omega_2+\omega_3)\frac{1}{1+H_1(\omega_1)B(\omega_1)}\} \\ &\cdot \{B(\omega_1+\omega_2)H_2(\omega_2,\omega_3)H_2(\omega_2,\omega_3)\frac{1}{1+H_1(\omega_2)B(\omega_2)}\} \\ &\cdot \frac{1}{1+H_1(\omega_3)B(\omega_3)}) \cdot (\frac{1}{1+H_1(\omega_2+\omega_3)B(\omega_2+\omega_3)})] \\ &\cdot \frac{1}{1+H_1(\omega_1+\omega_2+\omega_3)B(\omega_1+\omega_2+\omega_3)} \end{aligned} \quad (60)$$

Where $\tilde{G}_3$ is unsymmetrized. The application of feedback has reduced the gain of the linear transfer function by the factor $L(\omega) = 1+H_1(\omega)B(\omega)$. The second-order expression reveals that the application of feedback reduces the gain by the factor of $L(\omega_1)L(\omega_2)L(\omega_1+\omega_2)$. The third-order term indicates that the feedback has reduced its gain by $L(\omega_1)L(\omega_2)L(\omega_3)L(\omega_1+\omega_2+\omega_3)$ (and further terms at the frequencies of the two possible second order kernels).

From the above discussion, several observations can be made. First, the attenuation in each output term is related to its generating kernel-order; The higher order terms are reduced significantly, however, the reduction achieved for the lower order terms may not be sufficient for certain applications. This approach is only useful to systems in which feedback can be provided conveniently. Clearly,

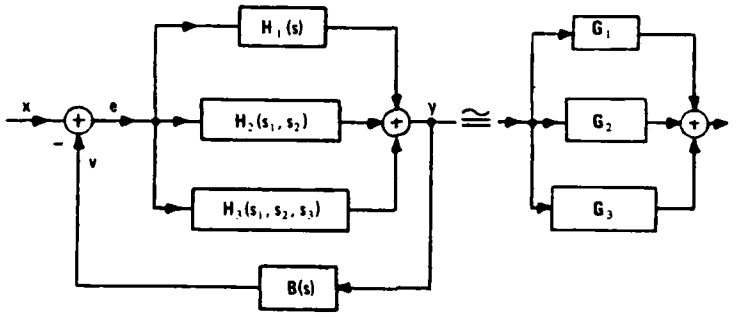


Fig. 6. Volterra model of a transistor with linear feedback.

feedback is not easily implemented if the nonlinearity of the entire link -- transmitter, channel and receiver -- is to be compensated. Finally, using linear feedback to reduce interference from the NLTFs also reduces the gain of the linear TF.

In our method of post-compensation (to be discussed later) these disadvantages are alleviated to a large degree.

CHAPTER III

INTERMODULATION AND CROSS-MODULATION IN NONLINEAR CHANNELS AND AMPLIFIERS

The phenomena of intermodulation and cross-modulation occur when two or more signals are mixed in a nonlinear element in such a way that multiplicative combinations of these signals are produced. New frequencies are generated (intermodulation), and intelligence from one carrier may jump on to another carrier (cross-modulation). This occurs in receiver input circuits, as well as in transmitters, and sometimes in a nonlinear element in the channel. As far as transmitters are concerned, the process involves the reception of an unwanted signal by the transmitting antenna, which conducts it back to the final stage of the power amplifier, where it is mixed with the transmitted signal. The process is therefore of greatest significance when both the unwanted signal and the nonlinear product are within the passband of the final amplifier. This can only occur if the nonlinear device has a characteristic of odd order. In this chapter several types of nonlinear distortion will be discussed.

Distortion Caused by Static Nonlinearities

Static nonlinearities may be characterized by a power series with constant coefficients. For example, a single input, single output static nonlinearity may be characterized as

$$y(t) = a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + \ldots + a_n x^n(t) + \ldots$$

$$= \sum_{n=1}^{\infty} a_n x^n(t). \qquad (61)$$

This type of nonlinearity is sometimes referred to as a memoryless nonlinearity, which is to say that the present output is only a function of the present input values and does not depend upon the past input values. A nonlinear resistor is an example of a static nonlinearity.

CROSS-MODULATION DISTORTION

Cross-modulation occurs when the modulation information of one carrier migrates to another carrier that did not have these sidebands. To illustrate the phenomenon that takes place in cross-modulation, consider a nonlinear device that has an input-output characteristic specified by a power series. Let the input to this device be the sum of an unmodulated carrier with frequency Ω_2, and a carrier with frequency Ω_1, modulated with a tone of frequency Ω_m:

$$x(t) = V(1+b_1\cos(\Omega_m t))\cos(\Omega_1 t) + V\cos(\Omega_2 t) \qquad (62)$$

The spectrum of this signal is shown in Fig. 7. Let us assume the terms of the static nonlinearity have negligible contribution beyond the cubic term. The terms of interest lie at the frequencies Ω_1 and Ω_2, these are the cross-modulated terms. The linear part of the response is

$$y_1(t) = a_1 V[(1+b_1\cos(\Omega_m t))\cos(\Omega_1 t)+\cos(\Omega_2 t)]. \qquad (63)$$

Using the identity

$$\cos(\alpha)\cos(\beta) = \frac{1}{2}\left[\cos(\alpha-\beta)+\cos(\alpha+\beta)\right]$$

(63) may be written as

$$y_1(t) = a_1 V\left[\cos(\Omega_1 t) + \frac{b_1}{2}\cos(\Omega_1-\Omega_m)t + \frac{b_1}{2}\cos(\Omega_1+\Omega_m)t + \cos(\Omega_2 t)\right] \quad (64)$$

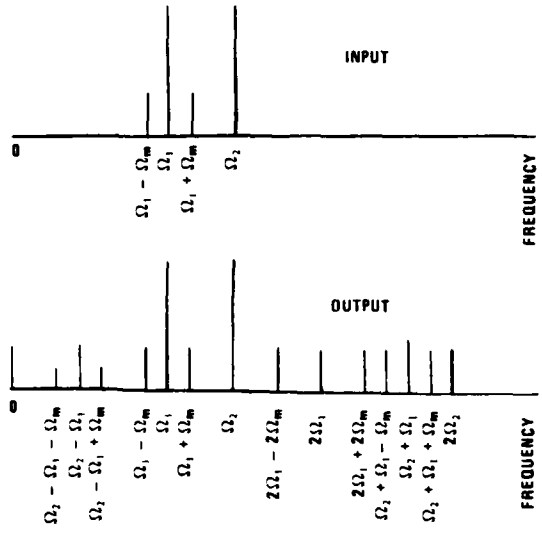


Fig. 7. Spectrum of quadratic nonlinearily generated cross-modulation.

The output spectrum is similar to (in fact proportional) to the input spectrum. The quadratic, or second-order, response is

$$\begin{aligned} y_2(t) = \frac{1}{4}\{a_2 V^2[&2(1+\cos(2\Omega_1 t))+2b_1(\cos(\Omega_m t)+\cos(2\Omega_1-\Omega_m)t) \\ &+ 2b_1(\cos(\Omega_m t)+\cos(2\Omega_1\pm\Omega_m)t)+2(\cos(\Omega_1+\Omega_2)t) \\ &+ \frac{b_1}{2}(4\cos(\Omega_1-\Omega_2)t)+b_1^2(1+\cos(2\Omega_1-2\Omega_m)t) \\ &+ b_1^2(\cos(2\Omega_m t)+\cos(2\Omega_1 t)+b_1(\cos(\Omega_2-\Omega_1+\Omega_m)t \\ &+ 2\cos(\Omega_2+\Omega_1-\Omega_m)t)+\frac{b_1^2}{2}(1+\cos(2\Omega_1+2\Omega_m)t) \\ &+ 2b_1(\cos(\Omega_2+\Omega_1\pm\Omega_m)t+(1+\cos(2\Omega_2 t))]\} \end{aligned} \quad (65)$$

Let us introduce a notation to reduce the length of the expressions. Let $\cos(\Omega)t$ be represented by C_Ω and $\cos(\Omega_1+\Omega_2)t$ by $C_{\Omega_1+\Omega_2}$. Then

the previous expression becomes

$$y_2(t) = \frac{1}{4}\{a_2V^2[2(1+\frac{b_1^2}{2})+4b_1C_{\Omega_m}+b_1^2C_{2\Omega_m}+2b_1(C_{\Omega_3-\Omega_m}+C_{\Omega_3+\Omega_m}$$

$$+ 2(C_{\Omega_3-\Omega_m}+C_{\Omega_4+\Omega_m})+b_1(C_{2\Omega_1-\Omega_m}+C_{2\Omega_1+\Omega_m})$$

$$+ 2(1+\frac{b_1^2}{2})C_{2\Omega_1}+2b_1(C_{\Omega_4-\Omega_m}+C_{\Omega_4+\Omega_m})$$

$$+ 2b_1(C_{2\Omega_1-2\Omega_m}+C_{2\Omega_1+2\Omega_m})+C_{2\Omega_2}]\} \qquad (66)$$

even more compactly,

$$y_2(t) = \frac{1}{4}\{a_2V^2[2(2+\frac{b_1^2}{2})+4b_1C_{\Omega_m}+b_1^2C_{2\Omega_m}+4C_{\Omega_3}+2b_1C_{\Omega_3\pm\Omega_m}$$

$$+ 2(1+\frac{b^2}{2})C_{2\Omega_1}+2b_1C_{2\Omega_1\pm\Omega_m}+4C_{\Omega_4}+2b_1C_{\Omega_4\pm\Omega_m}+2C_{2\Omega_2}]\} \qquad (67)$$

where $\Omega_3 = \Omega_2-\Omega_1$ and $\Omega_4 = \Omega_2+\Omega_1$.

Notice that there are no cross-modulation terms in this case.

The cubic or third-order term in the series requires a method for tabulating the frequency components. This will allow us to view the various terms faster and this approach can be used to analyze higher order nonlinearities. For the input

$$x(t) = V\{C_{\Omega_1}+\frac{b_1}{2}(C_{\Omega_1-\Omega_m}+C_{\Omega_1+\Omega_m})+C_{\Omega_2}\}$$

the cubic output is

$$y_3(t) = a_3x^3(t).$$

We may write the input in an expanded form, namely

$$x(t) = A_1x_1(t)+A_2x_2(t)+A_3x_3(t)+A_4x_4(t)$$

$$= \sum_{i=1}^{4} x_{\Omega_i}$$

where

$$x_{\Omega_i} \triangleq A_i \cos(\Omega_i t)$$

Then the output, $y_3(t)$ may be computed by a systematic combination of the input constituents as follows: the order is three therefore the number of columns (containing the input constituents) is also three. See Table 1. Each row represents a distinct combination of the three constituents from the available four sinusoids. For $y_3(t)$ the components of the output are listed in Table 1. The multipliers 1,3 or 6 in the right half of the table arise from the number of distinct permutations of the indices of the constituent sinusoids.

Using the cosine identities

$$\cos(\alpha)\cos(\beta)\cos(\lambda) = \frac{1}{4}\left(C_{\alpha-\beta+\lambda}+C_{\alpha+\beta-\lambda}+C_{\alpha-\beta-\lambda}+C_{\alpha+\beta+\lambda}\right)$$

$$\cos^2(\alpha)\cos(\beta) = \frac{1}{4}\left(C_{2\alpha-\beta}+2C_{\beta}+2C_{2\alpha+\beta}\right)$$

$$\cos^3(\alpha) = \frac{1}{4}\left(3C_{\alpha}+C_{3\alpha}\right)$$

The terms that produce cross-modulation are

$$3V^3 b_1 C_{\Omega_1} C_{\Omega_1-\Omega_m} C_{\Omega_2} + 3V^3 b_1 C_{\Omega_1+\Omega_m} C_{\Omega_2}$$

$$= \frac{3}{4} b_1 V^3 \{C_{\Omega_m-\Omega_2}+C_{\Omega_m+\Omega_2}+ \ldots +C_{-\Omega_m+\Omega_2}+C_{-\Omega_m-\Omega_2}\}$$

$$= \frac{3}{2} b_1 V^3 \{C_{\Omega_2-\Omega_m}+C_{\Omega_2+\Omega_m}\}+ \ldots \qquad (68)$$

Actually since

$$y_3 = a_3 x^3$$

$$= 3a_3 b_1 V^3 \cos(\Omega_m t)\cos(\Omega_2 t) + \ldots$$

TABLE 1
COMPUTATION OF OUTPUT FREQUENCIES GENERATED BY THIRD-ORDER NONLINEARITIES

permutations	translation
1 1 1	$A_1^3\cos^3(\Omega_1 t)$
1 1 2	$3A_1^2A_2\cos^2(\Omega_1 t)\cos(\Omega_2 t)$
1 1 3	$3A_1^2A_3\cos^2(\Omega_1 t)\cos(\Omega_3 t)$
1 1 4	$3A_1^2A_4\cos^2(\Omega_1 t)\cos(\Omega_4 t)$
1 2 2	$3A_1A_2^2\cos(\Omega_1 t)\cos^2(\Omega_2 t)$
1 2 3	$6A_1A_2A_3\cos(\Omega_1 t)\cos(\Omega_2 t)\cos(\Omega_3 t)$
1 2 4	$6A_1A_2A_4\cos(\Omega_1 t)\cos(\Omega_2 t)\cos(\Omega_4 t)$
1 3 3	$3A_1A_3^2\cos(\Omega_1 t)\cos^2(\Omega_3 t)$
1 3 4	$6A_1A_3A_4\cos(\Omega_1 t)\cos(\Omega_3 t)\cos(\Omega_4 t)$
1 4 4	$3A_1A_4^2\cos(\Omega_1 t)\cos^2(\Omega_4 t)$
2 2 2	$A_2^3\cos^3(\Omega_2 t)$
2 2 3	$3A_2^2A_3\cos^2(\Omega_2 t)\cos(\Omega_3 t)$
2 2 4	$3A_2^2A_4\cos^2(\Omega_2 t)\cos(\Omega_4 t)$
2 3 3	$3A_2A_3^2\cos(\Omega_2 t)\cos^2(\Omega_3 t)$
2 4 4	$3A_2A_4\cos(\Omega_2 t)\cos^2(\Omega_4 t)$
3 3 3	$A_3^3\cos^3(\Omega_3 t)$

"Table 1 (cont'd)"

3 3 4	$3A_3^2A_4\cos^2(\Omega_3 t)\cos(\Omega_4 t)$
3 4 4	$3A_3A_4^2\cos(\Omega_3 t)\cos^2(\Omega_4 t)$
4 4 4	$A_4^3\cos^3(\Omega_4 t)$

The carrier term may be approximated as

$$a_1 V\cos(\Omega_2 t) + \frac{3}{4} a_3 V^3\cos(\Omega_2 t)$$

$$\approx a_1 V\cos(\Omega_2 t) \qquad (69)$$

Thus, the cross-modulation at ω_2 is

$$y_{cm}(t) = a_1 V(1+b_{cm}\cos(\Omega_m t))\cos(\Omega_2 t)$$

$$b_{cm} = 3\frac{a_3}{a_1} b_1 V^2 \qquad (70)$$

The cross-modulation factor is the ratio of b_{cm} to b_1 is

$$cm_o = \frac{b_{cm}}{b_1}$$

$$= \frac{3a_3}{a_1} V^2 \qquad (71)$$

As an example consider $a_1 = 10$, $a_3 = 0.02$, $b_1 = 0.5$, $V = 1$ volt. Note that the amount of modulation is 50%. Then

$$b_{cm} = 3(0.002)(0.5)(1)^2$$

$$= 0.003$$

$$v_{cm}(t) = 10(1+ 0.003\cos(\Omega_m t))\cos(\Omega_2 t)$$

$$CM_o = \frac{b_{cm}}{b_1} = 0.006$$

This index measures the amount of modulation which has leaked onto the unmodulated carrier.

CROSS-MODULATION FOR NONLINEARITIES WITH MEMORY

Let the input be the sum of a continuous carrier and an amplitude modulated carrier as before

$$x(t) = V(1+b_1\cos\Omega_m t)\cos\Omega_1 t + V\cos\Omega_2 t$$

this can be written in a complex phasor form as

$$x(t) = \frac{V}{4}\left[2e^{j\Omega_1 t} + b_1 e^{j(\Omega_1-\Omega_m)t} + b_1 e^{j(\Omega_2-\Omega_m)t} + 2e^{j\Omega_2 t} \right.$$

$$\left. + 2e^{-j\Omega_1 t} + b_1 e^{-j(\Omega_1-\Omega_m)t} + b_1 e^{-j(\Omega_2-\Omega_m)t} + 2e^{-j\Omega_2 t} \right] \qquad (72)$$

To reduce complexity, let

$$e_\Omega = e^{j\Omega t}$$

Then we may write x(t) as

$$x(t) = \frac{V}{4}\left[2e_{\Omega_1} + b_1 e_{\Omega_1-\Omega_m} + b_1 e_{\Omega_1+\Omega_m} + 2e_{\Omega_2} \right] + CC \qquad (73)$$

where CC stands for complex conjugate terms.

If we use the tabulation technique shown earlier and select only the terms that contribute to the cross-modulation namely, the output component at the frequency $\Omega_2 \pm \Omega_m$ turns out to be

$$Y_{\Omega_2\pm\Omega_m} = \frac{3!}{16} V^3 b_1 \left[H_3(\Omega_1, -\Omega_1+\Omega_m, \Omega_2) e_{\Omega_2+\Omega_m} \right.$$

$$+ H_3(-\Omega_1, \Omega_1+\Omega_m, \Omega_2) e_{\Omega_2+\Omega_m} + H_3(\Omega_1, -\Omega_1-\Omega_m, \Omega_2) e_{\Omega_2-\Omega_m}$$

$$\left. + H_3(-\Omega_1, \Omega_1-\Omega_m, \Omega_2) e_{\Omega_2-\Omega_m} + CC \right] \qquad (74)$$

To discern the elements that cause cross-modulation, we need to express this equation differently. We will assume H_3 is a continuous function and that the frequency of modulation is much less than the carrier frequency,

$$\Omega_m << \Omega_1$$

We may now write the cross-modulation terms generated by the third-order nonlinearity as

$$Y_{\Omega_2 \pm \Omega_m} = \frac{3}{4} V^3 b_1 [H_3(\Omega_1, -\Omega_1, \Omega_2) e_{\Omega_2 + \Omega_m} + H_3(\Omega_1, -\Omega_1, \Omega_2) e_{\Omega_2 - \Omega_m} + CC]$$

$$= \frac{3}{2} b_1 V^3 |H_3(\Omega_1, -\Omega_1, \Omega_2)| [\cos((\Omega_2 - \Omega_m)t + \beta_2)$$

$$+ \cos((\Omega_2 + \Omega_m)t + \beta_2] \qquad (75)$$

where $\beta_2 = \text{phase}(H_3(\Omega_1, -\Omega_1, \Omega_2))$

If we look at the third-order system response to the carrier term

$$Y_3 = \frac{V}{2} H_1(\Omega_2) e_{\Omega_2} + \frac{V^3}{2} H_3(\Omega_2, \Omega_2, -\Omega_2) e_{\Omega_2} + CC$$

$$= \frac{V}{2} H_1(\Omega_2) e_{\Omega_2} + CC$$

$$= V |H_1(\Omega_2)| \cos(\Omega_2 t + \beta_1) \qquad (76)$$

where $\beta_1 = \text{phase}\{H_1(\Omega_2)\}$

Thus the $\Omega_2 t$ term becomes

$$y_{cm}(t) = V |H_1(\Omega_2)| \cos(\Omega_2 t + \beta_1)$$

$$+ 3 b_1 V^3 |H_3(\Omega_1, -\Omega_1, \Omega_2)| \cos(\Omega_2 t + \beta_2) \cos(\Omega_m t)$$

$$= |H_1(\Omega_1)| V [\cos(\Omega_2 t + \beta_1) + b_x \cos(\Omega_2 t + \beta_2) \cos(\Omega_m t)] \qquad (77)$$

where $b_x = 3b_1 \dfrac{|H_3(\Omega_1,-\Omega_1,\Omega_2)|}{|H_1(\Omega_2)|} V^2$

Therefore,

$$Y_{cm} = |H_1(\Omega_2)|V|[\cos(\Omega_2 t+\beta_1)+b_x\cos(\Omega_2 t+\beta_2)\cos\Omega_m t] \qquad (78)$$

The cross-modulation term is different from the memoryless case, there $y_{cm}(t)$ was pure amplitude modulation. Expressing the cross-modulation content as a function of β_1 and β_2 we must consider two cases:

Case I

$$\beta_1 = \beta_2$$

$$y_{cm}(t) = |H_1(\Omega_2)|V(1+b_x\cos\Omega_m t)\cos(\Omega_2 t+\beta_1)$$

This case generates pure amplitude modulation.

Case II

$$\beta_2 = \beta_1 + \frac{\pi}{2}$$

then the cross-modulation is

$$y_{cm}(t) = |H_1(\Omega_2)|V\cos(\Omega_2 t +\beta_1+b_x\cos\Omega_m t)$$

where $b_x << 1$

This case generates pure phase modulation.

INTERMODULATION DISTORTION

Intermodulation distortion is the process whereby beat signals at various sum and difference frequencies are produced. These signals are impossible to filter with a linear filter because, the interference frequencies lie within the passband of the uncorrupted signal. For example, consider an input consisting of the sum of two cosine waveforms,

$$x(t) = VC_{\Omega_1} + VC_{\Omega_2}$$

The intermodulation terms that would be generated by a second-order nonlinearity acting on this input are

$$Y_{2IM} = a_2 V^2 [C_{\Omega_2 - \Omega_1} + C_{\Omega_2 + \Omega_1}] \tag{79}$$

The third-order intermodulation products are at the frequencies $(2\Omega_1 \pm \Omega_2)$, $(2\Omega_2 \pm \Omega_1)$ and could lie within the passband of either carrier. The complete expression is found to be

$$Y_{3IM} = \frac{3}{4} a_3 V^3 [C_{2\Omega_1 \pm \Omega_2} + C_{2\Omega_2 \pm \Omega_1}] \tag{79}$$

The magnitude of these IM terms is

$$D_3 = \frac{3}{4} a_3 V^3 \tag{80}$$

Then the IM index is expressed by

$$IM_3 = \frac{D_3}{a_1} V = \frac{3a_3}{4a_1} V^2 \tag{81}$$

This concludes our discussion of interference types.

CHAPTER IV

COMPENSATION

Communication channels, both microwave and satellite, generally exhibit nonlinear characteristics. When frequency division multiplexing is employed, it is important to carefully assess the extent of possible intermodulation distortion. When this distortion exceeds permissible limits, it is reasonable to ask if it can be reduced by suitable compensation. In this chapter we present the compensation of both second and third-order NLTFs. The design methodology developed here has been implemented in FORTRAN programs VCOMP2 and VCOMP3 [12], [13]. Results of example runs on these programs are presented, which confirm the effectiveness of our technique in reduction of IM interference in a specified band.

The Pth-Order Inverse

Suppose the input-output relationship of Fig. 1, representing a weakly nonlinear system, is given by

$$\begin{aligned} y(t) &= H[x(t)] \\ &= H_1[x(t)]+H_2[x(t)]+H_3[x(t)]+ \cdots \end{aligned} \tag{82}$$

then a p-th order post-inverse is defined as follows. A nonlinear system G (also assumed to be representable in a Volterra series) is the p-th order post-inverse of the nonlinear system H if the composite system operator Q, given by,

$$Q[x(t)] = y(t)$$

$$= G[H[x(t)]] \tag{83}$$

contains only nonlinear terms of order p+1 or higher and its linear term is the identity mapping. That is

$$Q[x(t)] = x(t) + Q_{p+1}[x(t)] + \cdots \tag{84}$$

In this report our concern is limited to only second and third-order nonlinearities. We give below the conditions upon the NLTF's of H for the cases of second and third-order post-inverses: The following two conditions must be satisfied [1] for a second-order inverse to exist:

$$G_1H_1 = I \tag{85}$$

$$G_1H_2+G_2H_1 = 0 \tag{86}$$

For G to be the third-order inverse of H, three conditions must be satisfied, namely (85), (86) and the following

$$G_1H_3 + G_2\{H_1,H_2\} + G_3H_1$$

$$= G_1H_3+G_2[H_1+H_2]-G_2H_1-G_2H_2+G_3H_1 = 0 \tag{87}$$

Note that if $H_2 = 0$, i.e., the second-order nonlinearity is absent, then (87) reduces to

$$G_1H_3 +G_3H_1 = 0$$

Although the concept of a p-th order inverse is a useful one from a theoretical standpoint, its practical design utility is quite limited. In fact, even the first-order inverse defined formally by (85), has a somewhat limited utility in practice. Indeed, such an inverse would require the following relationship to be satisfied

$$G_1(s)H_1(s) = 1 \tag{88}$$

implying that $G_1(s)$ be the inverse of $H_1(s)$. This is impractical for two reasons. First, this would necessitate infinite gain for $G_1(s)$ at values of s where $H_1(s)$ is zero; second, if the transfer function $H_1(s)$ has zeros in the right-half plane, then $G_1(s)$ would contain poles in the right-half plane and would thus be unstable. In most communications applications, a global inverse is, in fact, not required. Indeed, if the signal spectrum is limited to the band $[\omega_1, \omega_2]$, then the objective should be to design G_1 so that G_1H_1 may appear to be almost equal to one (or, more generally, equal to $e^{-j\omega\tau}$) in this band.

As stated above, application to communication systems requires minimal transmission distortion only over a particular frequency band (or bands). Let us denote this band (or union of bands) as $\langle R_{roi} \rangle$; the subscript 'roi' signifies the region of interest. Also, let us denote the desired characteristic of the composite system as $C(\omega)$. Then, it is reasonable to state the desirable attributes of the compensator as follows: 1) The composite system Q, should have a characteristic as close to the desired as possible, i.e., G_1H_1 should approximate $C(\omega)$ over the band $\langle R_{roi} \rangle$, 2) the contributions of the NLTFs H_2, etc., in the band $\langle R_{roi} \rangle$ and possibly other interference regions of interest $\langle R_{IM} \rangle$ should be minimized, 3) the compensator should be stable, and 4) the structure of the compensator should be simple and relatively robust, i.e., insensitive to minor changes in the parameters of H_1, H_2, etc.

In the Appendix we present two simple examples dealing with the

compensation of linear distortion in bandpass systems. These examples are useful in that they provide insight into the structure of the compensator TF,(compensator pole location with respect to the system bandpass characteristic) needed to reduce linear distortion. This provides the basic compensator TF form to reduce the distortion caused by NLTFs in band-pass systems.

COMPENSATION OF THE QUADRATIC SYSTEM

A communication channel that exhibits only quadratic nonlinearity (and no higher order nonlinearity) may be represented by a linear transfer function in parallel with the quadratic TF, as shown in Fig. 8. This model might represent a transmitter, the channel, the front-end of a receiver, or a combination of these three.

When an input signal consisting of the sum of two sinusoids,

$$x(t) = V_1 \cos\Omega_1 t + V_2 \cos\Omega_2 t$$

is applied to the system of Fig. 8, the output is the sum of the linear and quadratic responses. The response of the linear block is

$$w_1(t) = V_1 |H_1(\Omega_1)| \cos((\Omega_1)t + \theta_{\Omega_1}) + V_2 |H_1(\Omega_2)| \cos((\Omega_2)t + \theta_{\Omega_2}) \quad (89)$$

where

$$\theta_{\Omega_i} = \text{Phase}\{H_1(\Omega_i)\}$$

The response of the quadratic block is

$$w_2(t) = \frac{1}{2}\left[V_1^2 |H_2(\Omega_1,\Omega_1)| \cos((2\Omega_1)t + \phi_1) + V_2^2 |H_2(\Omega_2,\Omega_2)| \cos((2\Omega_2 t + \phi_2)\right.$$
$$\left. + 2V_1 V_2 |H_2(\Omega_1,-\Omega_2)| \cos((\Omega_1-\Omega_2)t + \phi_3)\right.$$

$$+2V_1V_2|H_2(\Omega_1,\Omega_2)|\cos((\Omega_1+\Omega_2)t+\phi_4)$$

$$+2V_1^2H_2(\Omega_1,-\Omega_1)+2V_2^2H_2(\Omega_2,-\Omega_2)] \qquad (90)$$

where

$$\phi_1 = \text{Phase}\{H_2(\Omega_1,\Omega_1)\}$$

$$\phi_2 = \text{Phase}\{H_2(\Omega_2,\Omega_2)\}$$

$$\phi_3 = \text{Phase}\{H_2(\Omega_1,-\Omega_2)\}$$

$$\phi_4 = \text{Phase}\{H_2(\Omega_1,\Omega_2)\}$$

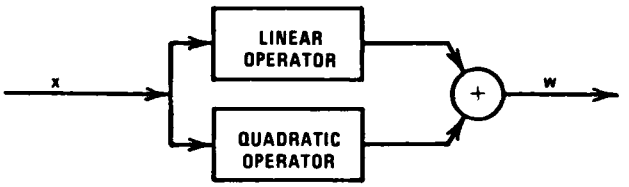


Fig. 8. Nonlinear communications channel model.

The system output then contains components with the following frequencies: Ω_1, Ω_2, $2\Omega_1$, $2\Omega_2$, $\Omega_1+\Omega_2$, $\Omega_1-\Omega_2$ and 0.

For reasons that will shortly become apparent, consider that Ω_1 and Ω_2 are nearly equal. Thus as shown in Fig. 9 the intermodulation terms occur near the baseband and around the frequency $2\Omega_1$. In realistic

situations the input spectrum is generally distributed over a frequency band with a bandwidth of Γ rad./s. This type of input might represent a communications signature. The regions of the output spectrum where IM distortion occurs may be obtained by substituting the boundary values of the input frequency band for Ω_1 and Ω_2 from the previous discussion,

$$\Omega_1 = \Omega_c + \frac{\Gamma}{2} \quad , \quad \Omega_2 = \Omega_c - \frac{\Gamma}{2}$$

it is readily seen that the output spectrum would now be as shown in Fig. 10.

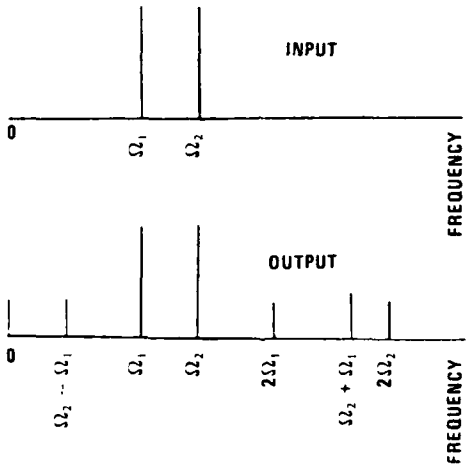


Fig. 9. Spectra of the input and output of the quadratic nonlinear system.

Note that the output interference energy is spread over two separate bands (one near the zero frequency and the other at double frequency).

A communications system may be compensated by pre-distorting the input signals with a nonlinearity which is an approximate 'inverse' of

the channel nonlinearities. Alternatively, one may use post-compensation wherein the compensator is placed after all of the system nonlinearities as shown in Fig. 11, for example, after the front-end of a receiver. Bell Laboratories have used pre-distortion for compensation on a narrow-band (compared to the carrier frequency) signal [6]. Their compensator structure as well as the theory developed are, however, not adequate for more general applications. We will use a post-compensation

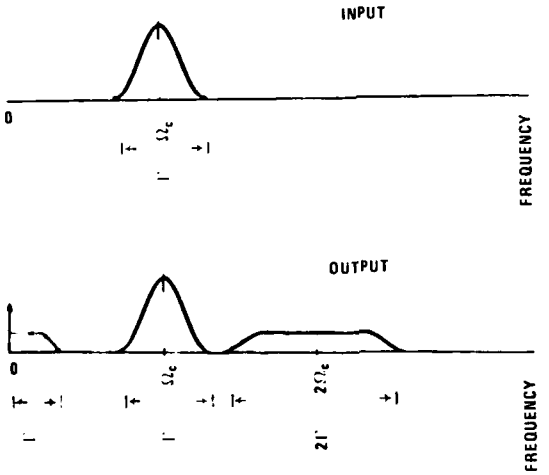


Fig. 10. Spectrum of the multiple frequency input to the quadratic nonlinear system.

strategy and will apply Volterra theory to develop a general approach for compensator design. With minor modifications the methodology can also be extended to pre-compensator design.

Post-Compensator Representation

The post-compensator transfer characteristic should have sufficient latitude so that by proper parameter selection reduction of the

nonlinear effects is possible. The post-compensator structure shown in Fig. 12 appeared promising and was proven effective in our studies. The sections of this compensator, are a gain constant for the linear kernel and a parametric nonlinear block. Only the nonlinear response of the **System Under Compensation** (SUC) will be compensated; the linear response will be left essentially unaltered. The output frequencies generated by the channel NLTF are out-of-band interference terms. Our compensator will reduce the interference at the frequency $2\Omega_1$. Having chosen the interference band of interest, the compensator design will now be discussed. The output of the compensator is

$$y = y_1 + y_2$$

where the output of the linear TF of the compensator is

$$y_1 = G_1[w_1] + G_1[w_2] + \underline{\cdots} \tag{91}$$

The compensators quadratic output is

$$y_2 = G_2[w_1] + \underline{G_2[w_2] + 2G_2\{w_1, w_2\} + \cdots} \tag{92}$$

where $G_2\{w_1, w_2\}$ is the bilinear Volterra operator [1].

Since our analysis is restricted, for the moment, to the second-order response, we will omit all terms shown underscored in (91) and (92).

Now, substituting for w_1 and w_2 from (89) and (90) and recognizing that $G_1(s) = K$,

$$\begin{aligned} y_1 = K\big[&V_1|H_1(\Omega_1)|\cos((\Omega_1)t + \theta_{\Omega_1}) + V_2|H_1(\Omega_2)|\cos((\Omega_2)t + \theta_{\Omega_2}) \\ &+ \tfrac{1}{2}\,[V_1^2|H_2(\Omega_1,\Omega_1)|\cos((2\Omega_1)t + \phi_1) + V_2^2|H_2(\Omega_2,\Omega_2)|\cos((2\Omega_2 t + \phi_2) \\ &+ 2V_1V_2|H_2(\Omega_1,-\Omega_2)|\cos((\Omega_1-\Omega_2)t + \phi_3) + 2V_1V_2|H_2(\Omega_1,\Omega_2)| \\ &\cdot \cos((\Omega_1+\Omega_2)t + \phi_4) + 2V_1^2H_2(\Omega_1,-\Omega_1) + 2V_2^2H_2(\Omega_2,-\Omega_2)]\big] \end{aligned} \tag{93}$$

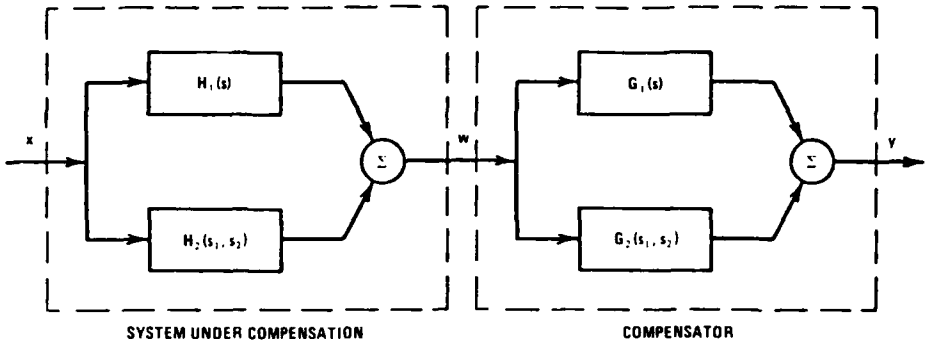


Fig. 11. Post-compensation of a communications system.

with quadratic nonlinearity.

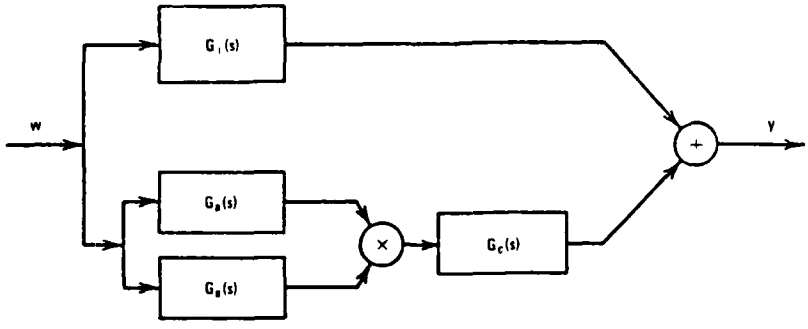


Fig. 12. Quadratic post-compensator structure.

where K is a scalar gain. The quadratic response is given by

$$y_2 = \frac{1}{2}\,[V_1^4|H_1(\Omega_1)|^2|G_2(\Omega_1,\Omega_1)|\cos((2\Omega_1 t)+2\theta_{\Omega_1}+\phi_1)$$

$$+V_2^4|H_1(\Omega_2)|^2|G_2(\Omega_2,\Omega_2)|\cos((2\Omega_2 t)+2\theta_{\Omega_2}+\phi_2)$$

$$+2[V_1V_2]^2|H_1(\Omega_1)H_1(\Omega_2)G_2(\Omega_1,-\Omega_2)|\cos((\Omega_1-\Omega_2)t+\theta_{\Omega_1}-\theta_{\Omega_2}+\phi_3)$$

$$+2[V_1V_2]^2|H_1(\Omega_1)H_1(\Omega_2)G_2(\Omega_1,\Omega_2)|\cos((\Omega_1+\Omega_2)t+\theta_{\Omega_1}+\theta_{\Omega_2}+\phi_4)$$

$$+2V_1^4H_1^2(\Omega_1)G_2(\Omega_1,-\Omega_1)+2V_2^4H_1(\Omega_2)G_2(\Omega_2,-\Omega_2)] \qquad (94)$$

Note that the right-hand side of (94) represents $G_2[w_1]$ since all other terms in (92) were omitted. From the earlier section we know that removal of the second-order output effects requires

$$G_1H_1 = I$$

$$G_1H_2+G_2H_1 = 0$$

or, in terms of the composite system TFs

$$Q_1 = I$$

$$Q_2 = 0$$

Since we stated that our goal is to reduce the distortion at $2\Omega_1$, we do so by setting terms involving this frequency to zero, i.e.,

$$\frac{V_1^2}{2}\ \text{Real}\{KH_2(\Omega_1,\Omega_1)e^{j2\Omega_1 t}+V_1^2H_1^2(\Omega_1)G_2(\Omega_1,\Omega_1)e^{j2\Omega_1 t}\} = 0 \qquad (95)$$

The above equation describes the relationship between H_1, H_2, K and G_2 to achieve reduction of interference at the frequency $2\Omega_1$. Since we have chosen $2\Omega_1$ as the frequency for compensation, any reduction of the interference at the frequency $2\Omega_2$ would be incidental. If reduction

of the IM distortion at both $2\Omega_1$ and $2\Omega_2$ were required, then the problem would have to be formulated in an entirely different manner; perhaps, the center frequency of the compensator along with its bandwidth could be appropriately chosen.

Multiple Frequency Compensation

We have stated an equation which describes the relationship between the SUC and the compensator to reduce intermodulation interference at $2\Omega_1$, for an input consisting of the sum of two unmodulated tones. Now let us compensate the IM interference for the case of an input which has a spectral bandwidth Γ. Now the reduction should be achieved over a band over which IM interference occurs:

$2\Omega_1 \pm \Gamma$, instead of a single frequency $2\Omega_1$.

Quadratic Mean-Square IM Criterion and a Related Cost Function

An index determining the amount of interference reduction is discussed next. It is natural to define a distortion energy criterion as

$$\varepsilon = \frac{1}{2\pi} \int_{-\infty}^{\infty} |Y_2(\omega)|^2 d\omega \tag{96}$$

The actual input signal x(t) is often not known, except that it is bandpass and real-valued (as is also the linear transfer function $H_1(s)$). For simplicity, we take the spectrum of the input to be uniform over this band, so that $Y_2(\omega_1,\omega_2) = Q_2(\omega_1,\omega_2)$. Now since

$$Q_2(\omega) \triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} Q_2(\omega-\psi_1,\psi_1)d\psi_1 \tag{97}$$

therefore our practical distortion criterion becomes

$$\varepsilon = \frac{1}{2\pi} \int_{\omega_0-\Delta}^{\omega_0+\Delta} |Q_2(\omega)|^2 d\omega \qquad (98)$$

where $\omega_0 \pm \Delta$ is the frequency range of interest.(i.e. The region in which interference is to be minimized.) It will be seen later that $\omega_0 = 2\omega_c$ and $\Delta = \Gamma$ if the IM interference at the double frequency is of concern.

The cost function must now be defined over the appropriate band of frequencies. Using equation (95) the intermodulation at the (ω_1, ω_2) frequency pair (in the associated-response frequency-plane) is

$$Q_2(\omega_1, \omega_2) = KH_2(\omega_1, \omega_2) + H_1(\omega_1)H_1(\omega_2)G(\omega_1, \omega_2) \qquad (99)$$

where the frequencies ω_1 and ω_2 each vary from $\Omega_1 - \frac{\Gamma}{2}$ to $\Omega_1 + \frac{\Gamma}{2}$

Computation of the Optimal Parameters

Parameter optimization applied to the design of dynamic systems involves the selection of a parameter set so as to enable the system to attain the best possible performance. We have selected for compensator design the nonlinear least squares minimization algorithm by Dennis et al [7]. This algorithm, in the large residual case, is more reliable than Gauss-Newton or Levenberg-Marquardt methods [8] and more efficient than the secant or variable metric methods algorithms [9], such as the Davidon-Fletcher-Powell method [10,11]. The algorithm amounts to a variation on Newton's method in which part of the Hessian matrix is computed exactly and part is approximated by a quasi-Newton updating method. To promote convergence from poor starting guesses, the algorithm uses a model/trust-region technique along with an adaptive choice of the model Hessian. The residual vector needed by NL2SNO is

constructed by converting the interference power density $|Q_2(\omega_1,\omega_2)|^2$ into the proper form. The transformation of the interference density in the ω_1,ω_2 plane into the real world frequency axis ω is done by searching the ω_1,ω_2 plane along constant sum (or difference) frequencies of the argument, and summing it. This is expressed by

$$Q_{IM}(\omega) = \int_{\omega_1+\omega_2=\omega} Q_2(\omega_1,\omega_2) \quad d\ell \tag{100}$$

where the integral is a line integral along the line defined by $\omega_1+\omega_2=\omega$. Note that $d\ell$ and $d\omega$ are orthogonal to each other. The ω_1,ω_2 plane along with the incremental line element $d\ell$ is shown in Fig. 13. If the band of concern were the baseband, then we would use the line defined by $\omega_1-\omega_2=\omega$ for the above integration.

In the computer program, the cost function is simply defined as

$$\epsilon_{PGM} = \sum_{i=-N}^{N} \sum_{k=-N}^{N} |Q_2(\omega_{1i},\omega_{2k})|^2 \tag{101}$$

where ω_{1i} and ω_{2k} are the grid points in the (ω_1,ω_2) frequency plane covering the square $[\Omega_1-\frac{\Gamma}{2}\ ,\ \Omega_1+\frac{\Gamma}{2}\ ]\ x\ [\Omega_2-\frac{\Gamma}{2}\ ,\ \Omega_2+\frac{\Gamma}{2}\ ]$. The center of the square grid is (Ω_1,Ω_2). The program NL2SNO uses

$$r_{i,k} \triangleq |Q_2(\omega_{1i},\omega_{2k})| \tag{102}$$

as the residual. The dimensionality of the residual vector is $(2N+1)^2$.

Initial Parameter Selection

In most optimization methods the a priori values affects the convergence, as well as the final solution, as local minima may exist

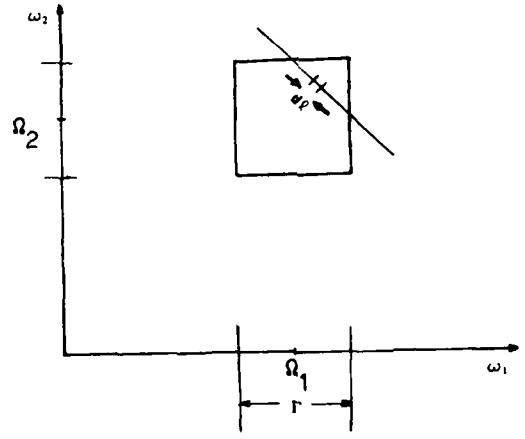


Fig. 13. The ω_1, ω_2 plane.

throughout the range of the cost function. Therefore a method of determining suitable initial parameter values (even in a semi-scientific way) is highly desirable. The approach we use is to compute the parameter values needed to compensate the SUC just at the center frequency (of the interference band). Then equation (99) becomes

$$0 = H_1^2(\Omega_c)G_2(\Omega_c,\Omega_c)+H_2(\Omega_c,\Omega_c) \tag{103}$$

where $\Omega_c = \Omega_1 = \Omega_2$. Rearranging, we obtain

$$G_2(\Omega_c,\Omega_c) = - \frac{H_2(\Omega_c,\Omega_c)}{H_1^2(\Omega_c)}$$

$$= x_H + jy_H = H_{MAG}\, e^{j\theta_H} \tag{104}$$

Shown in Fig. 14 is the magnitude of $Q_2(\Omega_1,\Omega_2)$ in the ω_1, ω_2 plane.

Consider a compensator which consists of a linear path with gain K, and a quadratic path as shown in Fig. 15. The transfer functions of the

blocks are

$$G_a(s) = \frac{s + b_1}{s + a_1} \quad , \; G_c(s) = \frac{b_0}{s + a_2} \tag{105}$$

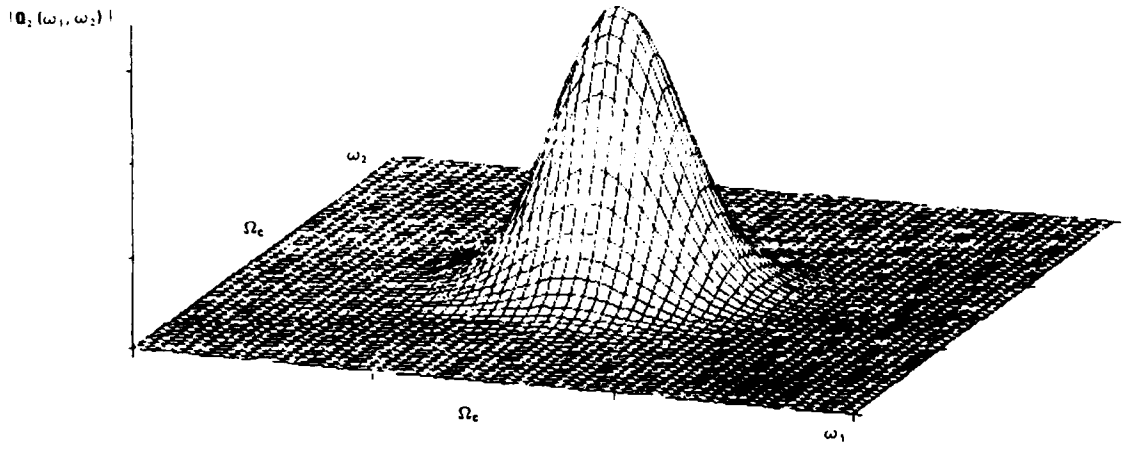


Fig. 14. Second-order distortion space.

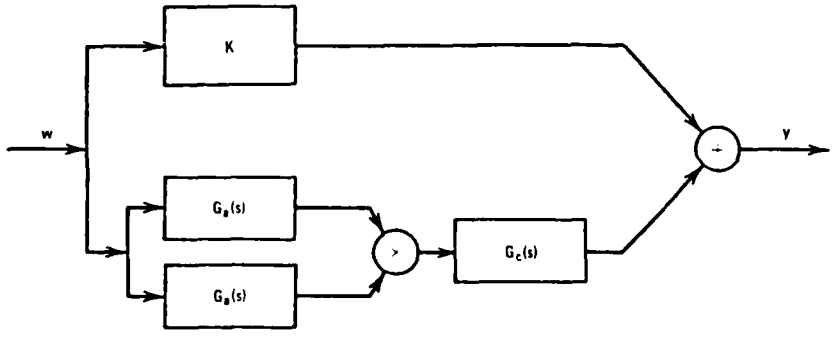


Fig. 15. Quadratic compensator structure

From this the NLTF of the quadratic path becomes

$$G_2(s_1,s_2) = \frac{b_0(s_1+b_1)(s_2+b_1)}{(s_1+a_1)(s_2+a_1)(s_1+s_2+a_2)} \qquad (105)$$

We shall call this a Type I Quadratic Compensator.

Suppose now we restrict our attention to a single unmodulated carrier with frequency Ω_c and the interference generated at the frequency $2\Omega_c$. Setting $s_1 = s_2 = j\Omega_c$ gives

$$G_2(\Omega_c,\Omega_c) = \frac{b_0(j\Omega_c+b_1)^2}{(j\Omega_c+a_1)^2(j2\Omega_c+a_2)} \qquad (106)$$

To compute the initial parameter vector, we equate the right-hand side of (104) and (105). We will now compute a_1, a_2 and b_1 so that the phase angle of the compensator is equal to θ_H. *In order* to select θ_G so that it is the additive inverse of θ_H, let

$$\theta_G = \theta_H - 2\pi, \quad \text{if } \theta_H > 0$$

$$\theta_G = \theta_H, \quad \text{if } \theta_H < 0$$

Then, equally distribute θ_G among the compensator poles at $-a_1$ and $\frac{-a_2}{2}$, and zero at $-b_1$ by defining,

$$\theta_D = \frac{\theta_G}{3}$$

Since θ_G is a negative angle, θ_D is positive.

The initial parameter values are then obtained by setting (the gain) b_0 of the compensator to 1 and solving expression (106) for each parameter value,

$$b_1 = -\frac{\Omega_c}{\tan(\frac{\theta_D}{2})} \quad , \quad a_1 = \frac{\Omega_c}{\tan(\frac{\theta_D}{2})}$$

$$a_2 = \frac{2\Omega_c}{\tan(\theta_D)} \tag{107}$$

Using these computed parameter values evaluate the compensator,

$$G_2(\Omega_c,\Omega_c) = G_{MAG}\, e^{j\theta_H}$$

setting,

$$b_0 = \frac{H_{MAG}}{G_{MAG}}$$

to adjust the gain parameter b_0 from its previously defined value of 1. This initial parameter method provides a starting location for compensation over the desired bandwidth. We will demonstrate the use of this result with an example.

Let

$$\Omega_c = 5 \text{ rad/sec}$$

$$H_1(\Omega_c) = 0.01 - j$$

$$H_2(\Omega_c,\Omega_c) = -0.09654 - j0.08098$$

Then from (103) the compensator response should be

$$G_2(\Omega_c,\Omega_c) = -0.0013 - j0.005$$

which gives

$$x_H = -0.0013 \ , \ y_H = -0.005$$

$$H_{MAG} = 0.00517 \ , \ \theta_H = -1.82516 \text{ rads}$$

and since θ_H is < 0 , $\theta_H = \theta_G$

$$\theta_D = 0.60839$$

$$b_1 = \frac{-5}{\tan(\ 0.30419)} = -15.9269$$

$$a_1 = \frac{5}{\tan(\ 0.30419)} = 15.9269$$

$$a_2 = \frac{10}{\tan(\ 0.60839)} = 164.1654$$

now compute the b_0 using these parameters,

$$2.2572 = \frac{H_{MAG}}{G_{MAG}}$$

The initial parameter values are then

$$\underline{p} = (2.2572\ -15.927\ 15.927\ 164.165)^T$$

EXAMPLES ON QUADRATIC COMPENSATION VIA PROGRAM VCOMP2

We now give two examples of quadratic interference reduction. Both examples have been generated by use of the computer program VCOMP2.

Example 1:

Consider a system with the Linear TF

$$H_1(s) = \frac{0.250025}{s^2 + 0.1s + 25.25}$$

which has a bandpass nature. The quadratic nonlinearity is of the form

$$H_2(s_1,s_2) = \frac{100}{(s_1+2)(s_2+2)(s_1+s_2+12)}$$

For the type of compensator to use we select the general form outlined in the previous section.

$$G_2(s_1,s_2) = \frac{b_0(s_1+b_1)(s_2+b_1)}{(s_1+a_1)(s_2+a_1)(s_1+s_2+a_2)}$$

The input is a narrowband signal with a center frequency of Ω_c= 5Mrad/s and a bandwidth Γ, of 2.0 Krad/s. The sum-squared error (i.e., sum

squared of $r(\omega_1, \omega_2)$ of the system with no compensation over a 5 x 5 point frequency grid) is ε_{PGM}= 3.94727. A simpler 3 x 3 example of the frequency grid is shown in Fig. 16. Note that the resulting band of interference is $(2\Omega_c - \Gamma,\ 2\Omega_c + \Gamma)$. The initial parameter vector is

$$\underline{p} = [2.2572\ -8.817\ 8.817\ 5.981]^T$$

After optimization the compensator parameter values are

$$\underline{p} = [43.65\ -129.9\ 234.9\ 59.72]$$

A plot of the SUC interference function $20\text{Log}(|Q_{IM}(\omega)|)$ versus the compensated one is shown in Fig. 17. The system quadratic sum-squared error is ε_{PGM}= 0.081724, a vast improvement over the uncompensated case.

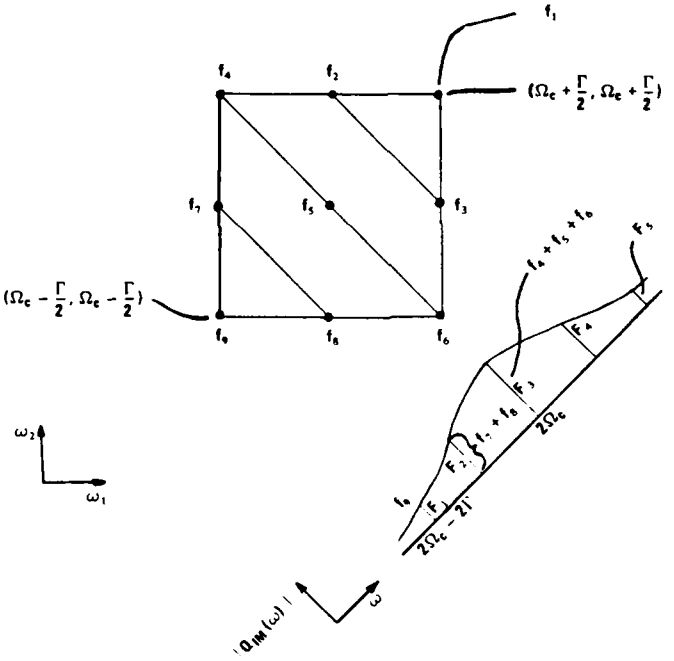


Fig. 16. 3 x 3 point frequency grid of $r(\omega_1, \omega_2)$

Example 2:

This example is similar to Example 1, except the amount of deviation is increased to 50.0 krad/s. The initial uncompensated sum-square error of the nonlinear component again over a 5 x 5 point frequency grid for $r(\omega_1, \omega_2)$ is ε_{PGM}= 8.301. The same initial parameter values as in the

last example are used. The compensated system error is 4.166. The final parameter values are

$$\underline{p} = [86.05\ -4533\ 6480\ 190]^T$$

In Fig. 18 a plot is given of the uncompensated and compensated

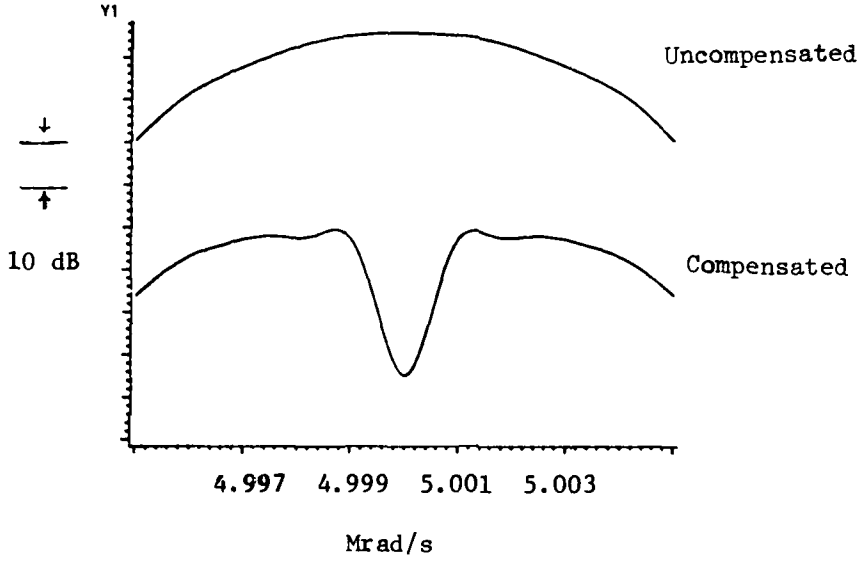


Fig. 17. Compensated versus uncompensated distortion.

interference function $20\text{Log}(|Q_{IM}(\omega)|)$ for this example. In the above examples we used a compensator whose blocks were single-pole blocks. A more effective reduction of the IM interference could be achieved with a more complex conpensator whose block TFs are second-order. Its NLTF is given by.

$$G_2(s_1,s_2) = \frac{b_0(s_1+b_1)(s_2+b_2)}{(s_1^2+a_2s_1+a_1)(s_2^2+a_2s_2+a_1)((s_1+s_2)^2+a_2(s_1+s_2)+a_1)} \quad (108)$$

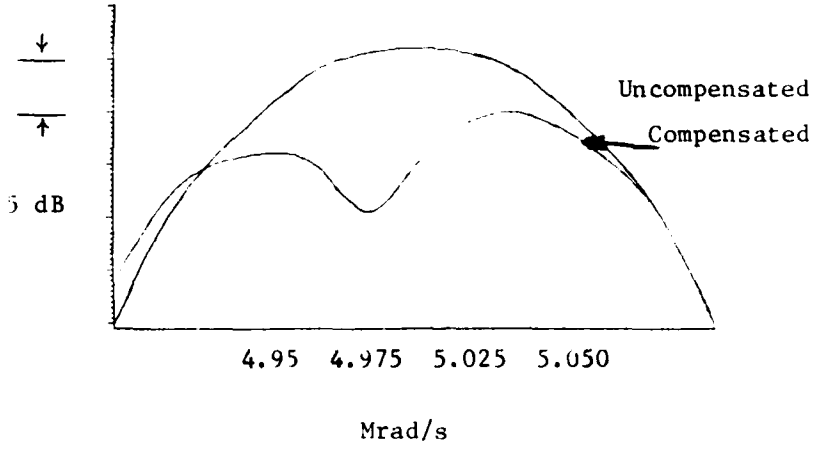


Fig. 18. Compensation results of the second example

This form of compensator will be called Type II Quadratic Compensator.

In this section we have explored the compensation of interference produced by a quadratic nonlinearity. As stated earlier, the interference generated by quadratic nonlinearities does not fall in the system's passband. However, the intermodulation terms generated by cubic nonlinearities, can fall within the main passband of the input to the SUC. Compensation of third-order nonlinearities is considered next.

COMPENSATION OF CUBIC NONLINEARITIES

In this section we present the design of the third-order nonlinear compensator using the configuration as shown in Fig. 19. The SUC here is shown in greater detail in Fig. 20. Note that its cubic nonlinearity may be represented mathematically by the NLTF

$$H_3(s_1,s_2,s_3) = H_a(s_1)H_a(s_2)H_a(s_3)H_d(s_1+s_2+s_3) \qquad (109)$$

Consider as an input to the SUC, a signal consisting of the sum of three sinusoids,

$$x(t) = V_1 \cos\Omega_1 t + V_2 \cos\Omega_2 t + V_3 \cos\Omega_3 t$$

the output of the SUC is the sum of the outputs pf the two paths; the response of the linear block is

$$\begin{aligned} w_1(t) = {} & V_1 |H_1(\Omega_1)| \cos((\Omega_1)t + \theta_{\Omega_1}) \\ & + V_2 |H_1(\Omega_2)| \cos((\Omega_2)t + \theta_{\Omega_2}) \\ & + V_3 |H_1(\Omega_3)| \cos((\Omega_3)t + \theta_{\Omega_3}) \end{aligned} \tag{110}$$

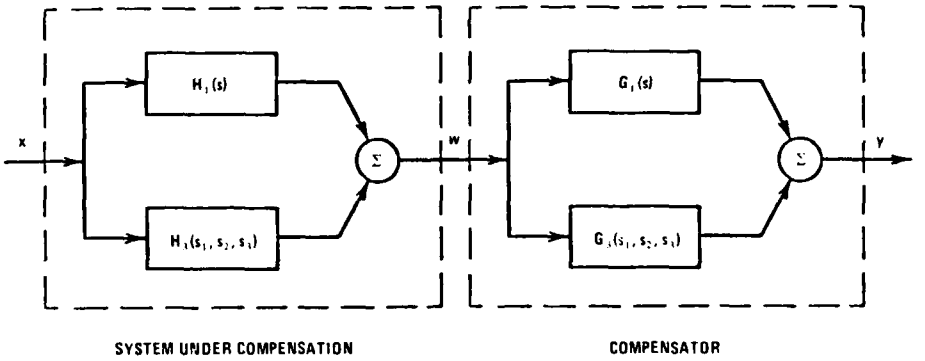


Fig. 19. Post-compensation of a communications system with cubic nonlinearity.

The response of the cubic block using expression (29) is

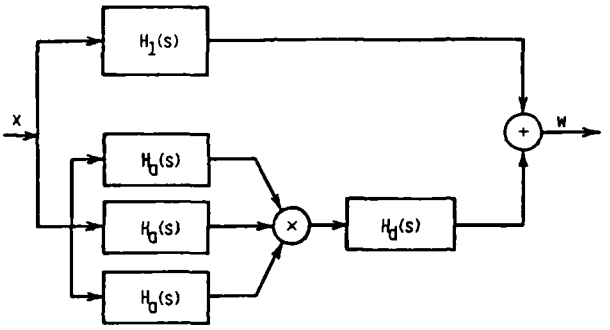


Fig. 20. Communications system model with cubic nonlinearity. (SUC)

$$w_3(t) = \frac{1}{4}\{\Gamma_{111}|Q_{111}|[3|H_a(\Omega_1)|\cos((\Omega_1)t + \theta_{\Omega_1})$$

$$+|H_d(3\Omega_1)|\cos((3\Omega_1)t + 3\theta_{\Omega_1} + \phi_{3\Omega_1})]$$

$$+\Gamma_{112}|Q_{112}|[2|H_a(\Omega_2)|\cos((\Omega_2)t + 2\theta_{\Omega_2})$$

$$+|H_d(2\Omega_1-\Omega_2)|\cos((2\Omega_1-\Omega_2)t + 2\theta_{\Omega_1} - \theta_{\Omega_2} + \phi_{2\Omega_1-\Omega_2})$$

$$+|H_d(2\Omega_1+\Omega_2)|\cos((2\Omega_1+\Omega_2)t + 2\theta_{\Omega_1} + \theta_{\Omega_2} + \phi_{2\Omega_1+\Omega_2})]$$

$$+\Gamma_{122}|Q_{122}|\,|[|H_d(\Omega_1-2\Omega_2)|\cos((\Omega_1-2\Omega_2)t + \theta_{\Omega_1} - 2\theta_{\Omega_2} + \phi_{\Omega_1-2\Omega_2})$$

$$+2|H_a(\Omega_1)|\cos((\Omega_1)t + 2\theta_{\Omega_1})$$

$$+|H_d(\Omega_1+2\Omega_2)|\cos((\Omega_1+2\Omega_2)t + \theta_{\Omega_1} + 2\theta_{\Omega_2} + \phi_{\Omega_1+2\Omega_2})]$$

$$+\Gamma_{222}|Q_{222}|[3|H_a(\Omega_2)|\cos((\Omega_2)t + \theta_{\Omega_2})$$

$$+|H_d(3\Omega_2)|\cos((3\Omega_2)t+3\theta_{\Omega_2}+\phi_{3\Omega_2}]$$

$$+\Gamma_{113}|Q_{113}|[2|H_a(\Omega_3)|\cos((\Omega_3)t+2\theta_{\Omega_3})$$

$$+|H_d(2\Omega_1+\Omega_3)|\cos((2\Omega_1+\Omega_3)t+2\theta_{\Omega_1}+\theta_{\Omega_3}+\phi_{2\Omega_1+\Omega_3})$$

$$+|H_d(2\Omega_1-\Omega_3)|\cos((2\Omega_1-\Omega_3)t+\theta_{2\Omega_1}-\theta_{\Omega_3}+\phi_{2\Omega_1-\Omega_3})]$$

$$+\Gamma_{123}|Q_{123}|[|H_d(\Omega_1-\Omega_2-\Omega_3)|\cos((\Omega_1-\Omega_2-\Omega_3)t+\theta_{\Omega_1}-\theta_{\Omega_2}-\theta_{\Omega_3}+\phi_{\Omega_1-\Omega_2-\Omega_3})$$

$$+|H_d(\Omega_1+\Omega_2-\Omega_3)|\cos((\Omega_1+\Omega_2-\Omega_3)t+\theta_{\Omega_1}+\theta_{\Omega_2}-\theta_{\Omega_3}+\phi_{\Omega_1+\Omega_2-\Omega_3})$$

$$+|H_d(\Omega_1-\Omega_2+\Omega_3)|\cos((\Omega_1-\Omega_2+\Omega_3)t+\theta_{\Omega_1}-\theta_{\Omega_2}+\theta_{\Omega_3}+\phi_{\Omega_1-\Omega_2+\Omega_3})$$

$$+|H_d(\Omega_1+\Omega_2+\Omega_3)|\cos((\Omega_1+\Omega_2+\Omega_3)t+\theta_{\Omega_1}+\theta_{\Omega_2}+\theta_{\Omega_3}+\phi_{\Omega_1+\Omega_2+\Omega_3})]$$

$$+\Gamma_{133}|Q_{133}|[H_d(\Omega_1-2\Omega_3)|\cos((\Omega_1-2\Omega_3)t+\theta_{\Omega_1}-2\theta_{\Omega_3}+\phi_{\Omega_1-2\Omega_3})$$

$$+2|H_a(\Omega_1)|\cos((\Omega_1)t+2\theta_{\Omega_1})$$

$$+|H_d(\Omega_1+2\Omega_3)|\cos((\Omega_1+2\Omega_3)t+\theta_{\Omega_1}+2\theta_{\Omega_3}+\phi_{\Omega_1+2\Omega_3})]$$

$$+\Gamma_{223}|Q_{223}|[2|H_a(\Omega_3)|\cos((\Omega_3)t+2\theta_{\Omega_3})$$

$$+|H_d(2\Omega_2-\Omega_3)|\cos((2\Omega_2-\Omega_3)t+2\theta_{\Omega_2}-\theta_{\Omega_3}+\phi_{2\Omega_2-\Omega_3})$$

$$+|H_d(2\Omega_2+\Omega_3)|\cos((2\Omega_2+\Omega_3)t+2\theta_{\Omega_2}+\theta_{\Omega_3}+\phi_{2\Omega_2+\Omega_3})]$$

$$+\Gamma_{233}|Q_{233}|[|H_d(\Omega_2+2\Omega_3)|\cos((\Omega_2+2\Omega_3)t+\theta_{\Omega_2}+2\theta_{\Omega_3}+\phi_{\Omega_2+2\Omega_3})$$

$$+2|H_a(\Omega_2)|\cos((\Omega_2)t+2\theta_{\Omega_2})$$

$$+|H_d(\Omega_2-2\Omega_3)|\cos((\Omega_2-2\Omega_3)t+\theta_{\Omega_2}-2\theta_{\Omega_3}+\phi_{\Omega_2-2\Omega_3})]$$

$$+\Gamma_{333}|Q_{333}|[3|H_a(\Omega_3)|\cos((\Omega_3)t+\theta_{\Omega_3})$$

$$+|H_d(3\Omega_3)|\cos((3\Omega_3)t + 3\theta_{\Omega_3} + \phi_{3\Omega_3})]\} \qquad (111)$$

The spectra of w_1 and w_3, the output of the SUC for an input consisting of the sum of three sinusoids is shown in Fig. 21. The in-band distortion falls at the following frequencies: $\Omega_1-\Omega_2+\Omega_3$, $\Omega_3-\Omega_1+\Omega_2$, $\Omega_2-\Omega_3+\Omega_1$. In the ensuing discussion we have chosen the distortion most adverse to our communication system to be at a band of frequencies around $\Omega_1-\Omega_2+\Omega_3$.

Post-Compensator Representation

The cubic compensator structure shown in Fig. 22 appeared promising and was tried first in our studies. The parallel paths of this compensator are a gain constant for the linear TF and a third-order nonlinear path. Only the nonlinear response of the SUC will be compensated; the linear response will be left essentially unaltered.

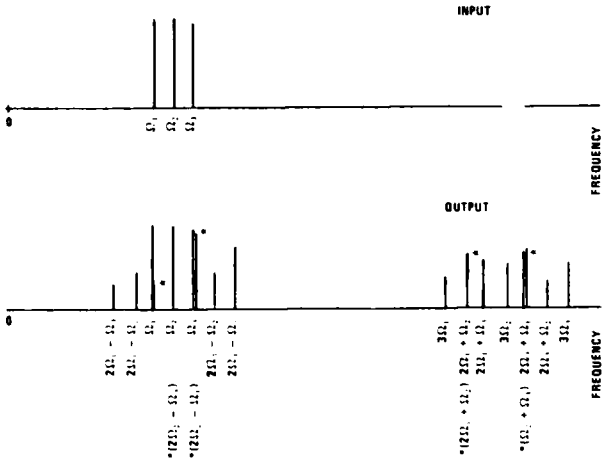


Fig. 21. Spectral plot of a third-order SUC output

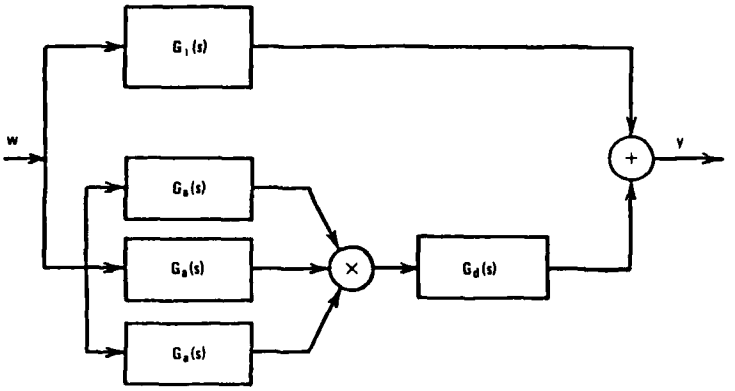


Fig. 22. Cubic nonlinearity compensator.

The output frequencies generated by the channel NLTF are out-of-band and in-band interference terms. Our choice of an interference frequency to be reduced is $\Omega_1-\Omega_2+\Omega_3$ which is an in-band term. Having chosen the interference frequency of interest, the compensator design will now be discussed. Note that at the moment we are dealing with the case of an input consisting of three sinusoids.

The output of the compensator is

$$y = y_1+y_3$$

where the output from the linear section of the compensator is

$$y_1 = KG_1[w_1]+KG_1[w_3] + \underline{\cdots} \tag{112}$$

$$y_1 = y_{1,L} + y_{1,N}$$

The compensators cubic output is

$$y_3 = G_3[w_1]+\underline{G_3[w_3]+ 3G_3\{w_1,w_1,w_3\}+ 3G_3\{w_1,w_3,w_3\} + \cdots} \tag{113}$$

Since our analysis is restricted, for the moment, to the third-order response, we will omit all terms shown underscored. Now, substituting

for w_1 and w_3 from (110) and (111) and retaining using only the distortion terms at the frequency of interest $\Omega_{IM} = \Omega_1 - \Omega_2 + \Omega_3$ we have

$$y_1\Big|_{\Omega_{IM}} = K6V_1V_2V_3|H_3(\Omega_1,-\Omega_2,\Omega_3)|\cos((\Omega_1-\Omega_2+\Omega_3)t + \phi_{\Omega_1-\Omega_2+\Omega_3})] \quad (114)$$

$$y_3\Big|_{\Omega_{IM}} = G_3[w_1]$$

$$= 6V_1V_2V_3|H_1(\Omega_1)H_1(-\Omega_2)H_1(\Omega_3)G_3(\Omega_1,-\Omega_2,\Omega_3)|$$

$$\cdot\cos((\Omega_1-\Omega_2+\Omega_3)t + \theta_{\Omega_1} + \theta_{-\Omega_2} + \theta_{\Omega_3} + \psi_{\Omega_1-\Omega_2+\Omega_3}) \quad (115)$$

where we have let G_1 be a scalar gain K.

From earlier discussion, we know that removal of the third-order output effects requires

$$G_1H_1 = I$$

$$G_1H_3 + G_3H_1 = 0$$

or, in terms of the composite system TFs,

$$Q_1 = I$$

$$Q_3 = 0$$

Since we stated that our goal is to reduce the IM interference at $\Omega_1 - \Omega_2 + \Omega_3$, we do so by setting the sum of the terms involving this frequency to zero, i.e.,

$$6KV_1V_2V_3|H_3(\Omega_1,-\Omega_2,\Omega_3)|\cos((\Omega_1-\Omega_2+\Omega_3)t + \phi_{\Omega_1-\Omega_2+\Omega_3})$$

$$+6V_1V_2V_3|H_1(\Omega_1)H_1(-\Omega_2)H_1(\Omega_3)G_3(\Omega_1,-\Omega_2,\Omega_3)|$$

$$\cdot\cos((\Omega_1-\Omega_2+\Omega_3)t + \theta_{\Omega_1} + \theta_{-\Omega_2} + \theta_{\Omega_3} + \psi_{\Omega_1-\Omega_2+\Omega_3}) = 0 \quad (116)$$

The above equation describes the relationship between H_1, H_3, K, and G_3

to achieve reduction of interference at the frequency $\Omega_{IM} = \Omega_1 - \Omega_2 + \Omega_3$.

Multiple Frequency Compensation

We have stated an equation which describes the relationship between the SUC and the compensator to reduce intermodulation interference at $\Omega_1 - \Omega_2 + \Omega_3$, for an input consisting of the sum of three unmodulated tones. Now let us consider the IM interference for the case of an input which has a spectral distribution over a bandwidth Γ. An example spectral distribution of the channel NLTF response is given in Figure 23. Now the reduction should be achieved over a band over which the IM interference occurs: $(\Omega_1 - \Omega_2 + \Omega_3) \pm \frac{3\Gamma}{2}$ instead of merely the frequency $\Omega_1 - \Omega_2 + \Omega_3$.

Cubic Distortion Criterion and a Related Cost Function

A cost function describing the amount of interference is formulated next. It is natural to define a distortion energy criterion as

$$\epsilon = \frac{1}{2\pi} \int_{-\infty}^{\infty} |Y_3(\omega)|^2 d\omega \tag{117}$$

The actual input signal x(t) is often not known, except that it is bandpass (as is also the linear transfer function $H_1(s)$). For simplicity, we take the spectrum of the input to be uniform over this band, so that $Y_3(\omega_1, \omega_2, \omega_3) = Q_3(\omega_1, \omega_2, \omega_3)$. Now if we define

$$Q_3(\omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int Q_3(\omega - \psi_1, \psi_1 - \psi_2, \psi_2) d\psi_1 d\psi_2 \tag{118}$$

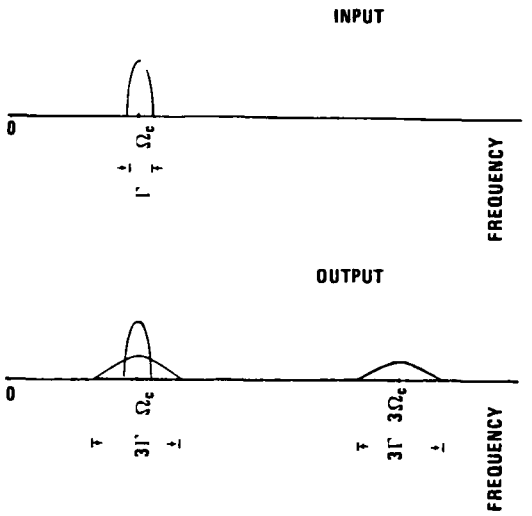


Fig. 23. Spectral plot of a third-order SUC output from an input consisting of a carrier Ω_c with bandwidth Γ.

then our practical distortion criterion becomes

$$\varepsilon = \frac{1}{2\pi} \int_{\omega_0-\Delta}^{\omega_0+\Delta} |Q_3(\omega)|^2 d\omega \qquad (119)$$

where $\omega_0 \pm \Delta$ is the frequency range of interest.(i.e. The region in which interference is to be minimized.) It will be seen later that $\omega_0 = \Omega_c$ and $\Delta = 3\Gamma$ where the IM interference at $\Omega_{IM} = \Omega_1 - \Omega_2 + \Omega_3$ is addressed.

The cost function must now be defined over the appropriate band of frequencies. Using equation (118) the intermodulation at the $(\omega_1, \omega_2, \omega_3)$ frequency coordinates (in the associated-response frequency-plane) is

$$Q_3(\omega_1,\omega_2,\omega_3) = KH_3(\omega_1,\omega_2,\omega_3)+H_1(\omega_1)H_1(-\omega_2)H_1(\omega_3)G(\omega_1,-\omega_2,\omega_3) \qquad (120)$$

where the frequencies ω_1, ω_2 and ω_3 each vary from $(\Omega_c - \frac{\Gamma}{2})$ to $(\Omega_c + \frac{\Gamma}{2})$. The residual vector needed by NL2SNO [4] is constructed by converting the interference density $|Q_3(\omega_1,\omega_2,\omega_3)|^2$ into the proper form. The transformation of the in-band interference power density in $(\omega_1,\omega_2,\omega_3)$ into the real world frequency ω is accomplished by integrating the values of Q_3 that lie on an interference surface $\sigma_{IM}(\omega)$ defined by

$$\omega_1 - \omega_2 + \omega_3 = \omega$$

This is expressed by

$$Q_{IM}(\omega) = \iint_{\omega=\omega_1-\omega_2+\omega_3} Q_3(\omega_1,\omega_2,\omega_3)\ d\sigma_{IM} \qquad (121)$$

where the integral is a surface integral along the plane defined by $\sigma_{IM}(\omega)$ for $\omega_1-\omega_2+\omega_3=\omega$. Note that $d\sigma_{IM}$ and $d\omega$ are orthogonal to each other.

In the computer program, the cost function is simply defined as

$$\varepsilon_{PGM} = \sum_{i=-N}^{N} \sum_{j=-N}^{N} \sum_{k=-N}^{N} |Q_3(\omega_{1i},\omega_{2j},\omega_{3k})|^2 \qquad (122)$$

where ω_{1i}, ω_{2j} and ω_{3k} are the grid points in the $(\omega_1,\omega_2,\omega_3)$ frequency space covering the cube $[\Omega_1 - \frac{\Gamma}{2}, \Omega_1 + \frac{\Gamma}{2}] \times [-\Omega_2 - \frac{\Gamma}{2}, -\Omega_2 + \frac{\Gamma}{2}] \times [\Omega_3 - \frac{\Gamma}{2}, \Omega_3 + \frac{\Gamma}{2}]$. The center of the cubic grid is $\Omega_1,-\Omega_2,\Omega_3$. The program NL2SNO uses

$$r_{i,j,k} \triangleq |Q_3(\omega_{1i},\omega_{2j},\omega_{3k})| \qquad (123)$$

as the residual. The dimensionality of the residual vector is $(2N+1)^3$. Shown in Fig. 24 is a plot of the $r_{i,j,k}$ grid. Fig. 25 shows the location of an example cube centered at $(\omega_1,\omega_2,\omega_3)$.

Fig. 26 shows one of the surfaces $\sigma_{IM}(\Omega_c)$ in the three-dimensional interference space for $\omega=\Omega_c$; this plane maps into the real world frequency Ω_c.

Delay Block in the Compensator Structure

Let us define the group delay of the contributions of the third-order intermodulation as

$$\tau_{IM}(\omega) \triangleq -\frac{\partial}{\partial\omega}\ \text{Phase}\{Q_{IM}(\omega)\} \tag{124}$$

We first considered the cubic compensator previously shown in Fig. 21 with nonlinear path NLTF given as

$$G_3(s_1,s_2,s_3) = \frac{b_0(s_1+b_1)(s_2+b_1)(s_3+b_1)}{(s_1+a_1)(s_2+a_1)(s_3+a_1)(s_1+s_2+s_3+c_2)} \tag{125}$$

which we will call the Type I compensator.(A fairly complete family listing of the useful cubic compensator structures is given in Table 2.) When the group delay of the third-order nonlinearity of the SUC was small, the optimized Type I compensator turns out to be stable (i.e., the block poles of the compensator turned out to be in the left-half s-plane). However, large values of SUC group delay τ_{IM} render the Type I compensator inadequate; for such cases the program VCOMP3 yields a) only a small reduction in ε_{PGM} or, b) unstable block poles for the compensator, or both.

TABLE 2

TABLE OF COMPENSATOR STRUCTURES

IDENTIFIER	G_1	G_a BLOCK TF IN G_3	G_d BLOCK TF IN G_3
Type I	K	$\frac{(s + b_1)}{(s + a_1)}$	$\frac{b_0}{(s + a_2)}$
Type I_D	The same as the Type I; except that $G_1 = Ke^{-s\tau}$		
Type IIa	K	$\frac{b_0 s^2 + b_1 s + b_3}{(s + 10\Omega_c)(s + 0.1\Omega_c)}$	$\frac{1}{(s + a_3)}$
Type IIa_D	The same as the Type IIa; except that $G_1 = Ke^{-s\tau}$		
Type IIb	K	$\frac{b_0 s^2 + b_1 s + b_3}{s^2 + a_1 s + a_0}$	$\frac{1}{(s + a_3)}$
Type IIb_D	The same as the Type IIb; except that $G_1 = Ke^{-s\tau}$		

Note: Structures I, I_D, IIb, and IIb_D are not guaranteed to be stable. However, as discussed in subsection 3.4, they can be reparametrized so that in the space of the new parameters these forms are guaranteed to be stable for every point.

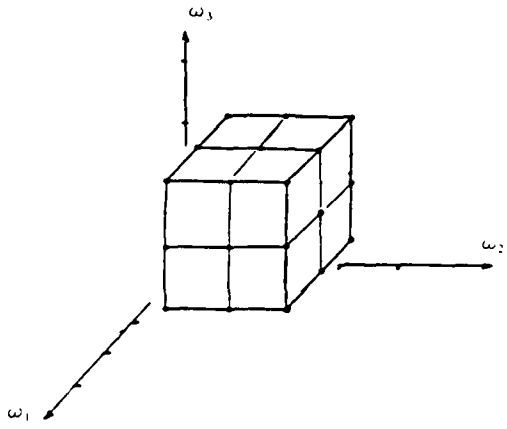


Fig. 24. $r_{i,j,k}$ grid

A modified compensator structure which can effectively reduce the IM interference is proposed next. This modified structure contains a delay block in the linear path of the compensator, as shown in Fig. 27. The compensator NLTF G_3 remains the same, thus the designation of this compensator is Type I_D. The incorporation of the delay shifts the phase of w_3 (SUC IM interference) so that the group delay of y_3 (the compensator cubic response to the input of w_1) duplicates the group delay of $y_{1,N}$(the delayed SUC IM interference). This altering of the w_3 group delay facilitates constructing a stable G_3 so that y_3 is the additive inverse of $y_{1,N}$. The delayed version of w_3 (i.e., $y_{1,N}$), has a phase characteristic that can be adequately cancelled by the response y_3

of the NLTF G_3 yet allowing the block poles to be in the left-half s-plane. Without the delay block, the compensator was able to reduce the IM interference only slightly, and had often returned

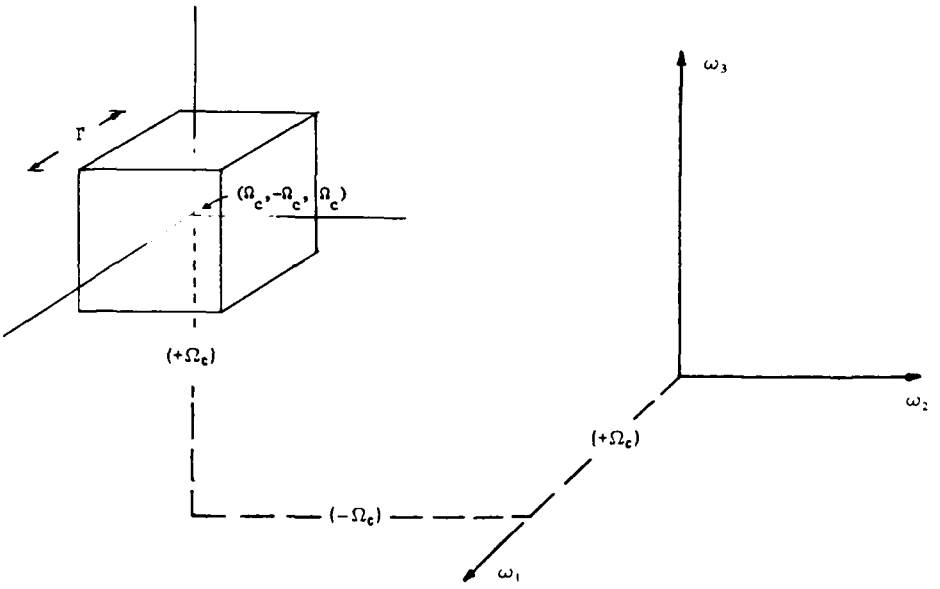


Fig. 25. Location of the in-band interference region in $Q_3(\omega_1, \omega_2, \omega_3)$

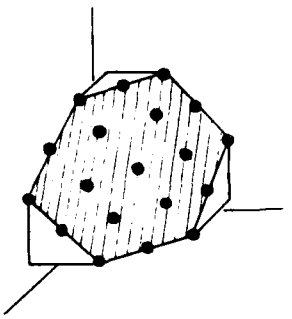


Fig. 26. $r_{i,j,k}$ grid with one interference surface shown.

unstable block poles. The output of the compensator is

$$y = y_{1,L} + y_{IM} + \underline{\cdots} \tag{126}$$

where $y_{1,L}$ is the linear response, and y_{IM} is the IM interference term, and the underscored terms are of order higher than three, hence omitted.

Note that

$$y_{1,L} = KD_{\tau}[w_1]$$

where D_{τ} is the delay operator, $D_{\tau}[w_1] = w_1(t-\tau)$ (127)

Clearly, the IM interference signal is

$$y_d(t) = KD_{\tau}[H_3[x(t)]] + G_3[w_1(t)]. \tag{128}$$

The effect of the group delay manifests into the composite system NLTF in the following manner:

$$Q_{(3)}(s_1,s_2,s_3) = Ke^{-(s_1+s_2+s_3)\tau} H_3(s_1,s_2,s_3) + H(s_1)H(s_2)H(s_3)G_3(s_1,s_2,s_3) \tag{129}$$

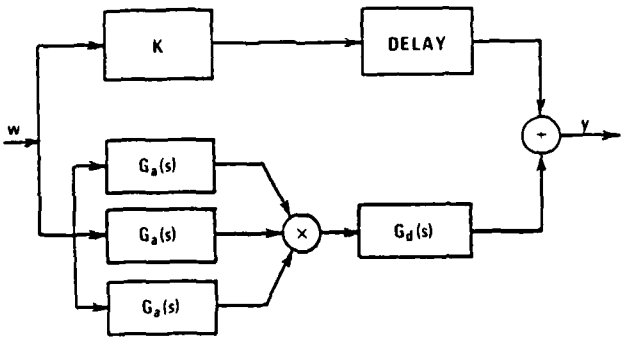


Fig. 27. Addition of a phase delay to the cubic compensator

The Type I_D cubic compensator provided an improvement over the Type I; the computer program VCOMP3 returns unstable block TFs only infrequently. Still when compensating an SUC with a large group delay it was occasionally unstable. This led to the development of yet another type of cubic compensator in which the poles of the block TFs are restricted to only stable values. In this compensator the form of the block TFs is the same as in the Type II; however the denominator parameters are now frozen as

$$a_0 = \Omega_c^2$$

$$a_1 = 0.1\Omega_c + 10\Omega_c$$

so that the the compensator poles are at $10\Omega_c$ and $0.1\Omega_c$. Thus,

$$G_a(s) = \frac{b_2 s^2 + b_1 s + b_0}{(s + (0.1)\Omega_c)(s + (10)\Omega_c)} \quad , \; G_d(s) = 1 \qquad (122)$$

This structure is called Type IIa cubic compensator. We now discuss a method to allow the compensator block poles to vary while simultaneously ensuring that they stay in the left half-plane.

Re-parmetrization of the Compensator Block TFs

To completely circumvent the possibility of obtaining unstable poles in the compensator block TFs, we re-parametrize the G_a and G_d block TFs in G_3. For brevity we will only consider Type IIb_d structure. The TF structure for $G_a(s)$ and $G_d(s)$ is expressed as:

$$G_a(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^2 + a_1 s + a_0}$$

$$G_d(s) = \frac{1}{s + c}$$

We now re-parametrize both transfer functions by use of the equations

$$\begin{aligned} b_0 &= p_1 \\ b_1 &= p_2 \\ b_2 &= p_3 \\ a_0 &= p_4^2 + \alpha \\ a_1 &= \gamma\sqrt{a_0} + p_5 \\ c &= p_6^2 \\ \tau &= p_7 \end{aligned} \tag{121}$$

The new parameter vector is

$$p = (p_1, p_2, p_3, p_4, p_5, p_6, p_7)$$

wherein only p_4, p_5 affect the poles of G_a, and p_6 affects the pole of G_d. Every point of this parameter space maps into a stable point in the $(b_0, b_1, b_2, a_1, a_0, c, \tau)$ space. Indeed it can be seen that the damping ratio of $G_a(s)$ is guaranteed to be greater than or equal to ζ. This form will be called the Type IIb_D. The poles of the Type IIb_D compensator are shown in Fig. 28. Notice that the compensator poles at s_1 and s_2 lie to the left of the contour Γ no matter what the values of the parameters p_4 and p_5 may be.

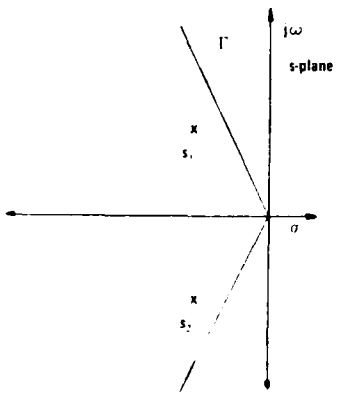


Fig. 28. Pole plot for modified Type IIb cubic compensator designs.

EXAMPLES OF INTERMODULATION REDUCTION USING THE PROGRAM VCOMP3

Results of example runs on the program VCOMP3 are presented, which confirm the effectiveness of our technique in reduction of IM

interference in a specified band.

Since most real world communication channels exhibit band-pass behavior, the model used for the channel is a Chebyshev bandpass filter of second, fourth and sixth-degree (passband ripple is 0.5 dB). The nonlinearity is third-order and produces a signal approximately -20 dB below the linear response. In the computer examples we vary the cutoff rate of the channel (degree), the NLTF of the channel. The compensator used is of Type II_D form shown earlier.

Compensation Examples by the Computer Program VCOMP3

We now give several numerical examples of cubic interference reduction by the computer program VCOMP3.

Example 1

Consider a system with the transfer function,

$$H_1(s) = \frac{0.571997(10^6)s}{s^2 + 0.571997(10^6)s + 24.9916(10^{12})}$$

which is a second-degree Chebyshev with a passband 4.9 Mrad/s to 5.1 Mrad/s. The system has a cubic nonlinearity of the form.

$$H_3(s_1, s_2, s_3) = \frac{4.6416}{(s_1 + 2)(s_2 + 2)(s_3 + 2)(s_1 + s_2 + s_3 + 12)}$$

This NLTF has a three dimensional impulse response resembling a lowpass filter. The type of compensator used is the stable cubic with linear delay, outlined in the previous section.

$$G_3(s_1, s_2, s_3, \underline{p})$$

$$= \frac{(p_2 s_1^2 + p_1 s_1 + p_0)(p_2 s_2^2 + p_1 s_2 + p_0)}{(s_1 + 0.1\omega_c)(s_1 + 10\omega_c)(s_2 + 0.1\omega_c)(s_2 + 10\omega_c)}$$

$$\cdot \frac{(p_2 s_3^2 + p_1 s_3 + p_0)}{(s_3 + 0.1\omega_c)(s_3 + 10\omega_c)}$$

with the linear section of the compensator defined as

$$G(s) = e^{-(s_1 + s_2 + s_3)p_4}$$

The input is a narrowband signal with a center frequency Ω_c of 5 Mrad/s and a bandwidth Γ of 200 Krad/s. The sum-squared error of the system with no compensation is $\varepsilon_{PGM} = 0.3036748$. The initial parameter vector is

$$\underline{p} = [-4.6416 \quad 1.587 \quad 10.33 \quad 0.0]^T$$

After optimization the compensator parameter values are

$$\underline{p} = [-839.3 \quad -32.18 \quad -33.21 \quad 1.265]$$

The poles of the compensator were fixed, only the gain and the location of the zeros were varied. We therefore know that this compensator is stable. The compensated system ε_{PGM} is 0.0003306 a vast improvement over the uncompensated case. A plot of the compensated nonlinear interference verses the uncompensated is given in Fig. 29.

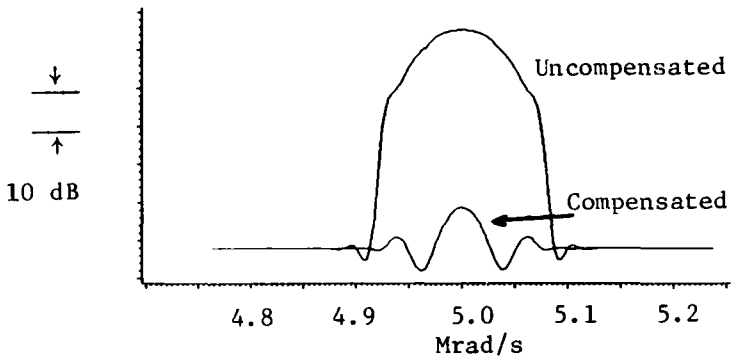


Fig. 29. Compensated versus the uncompensated interference.

Example 2

This example is similar to Example 1. The linear channel TF is unchanged; however, the blocks of the NLTF are all second-degree bandpass Chebyshev. Thus $H_a = H_d$ both have a somewhat narrow bandpass characteristic. The sum-squared error of the system with no compensation ε_{PGM} is 105.34013. The initial parameter vector is

$$\underline{p} = [-4.6416\ 0.0\ 0.0\ 0.0]^T$$

After optimization the compensator parameter values are

$$\underline{p} = [1647.3\ -85.75\ \ 66.65\ 1.579]$$

The compensated system ε_{PGM} is 0.0467988, again a vast improvement over the uncompensated case. A comparison of the compensated nonlinear interference and the uncompensated one is given in Fig. 30

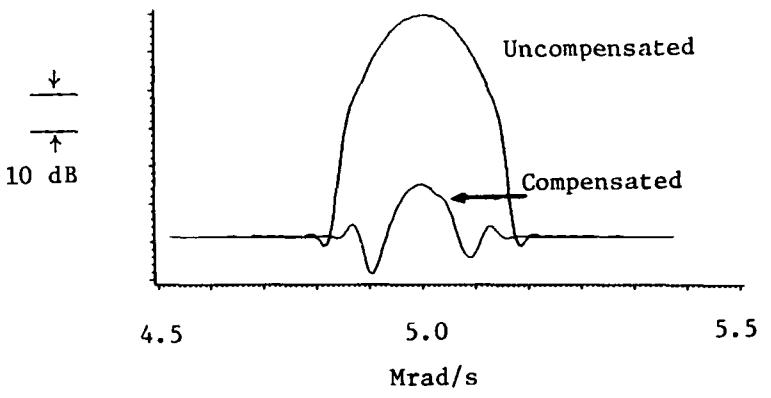


Fig. 30. Compensated versus the uncompensated interference.

Example 3

This example has a linear TF with a further reduced bandwidth. The linear characteristic of the channel is modeled by a fourth-degree Chebyshev with a passband of 4.9 Mrad/s to 5.1 Mrad/s.

$$H_1(s) = \frac{0.05714407s^2}{[s^4 + 0.28484732s^3 + 50.0437463s^2 + 7.11879269s + 624.580478]}$$

The sum-squared error of the system without compensation is ε_{PGM} =59.27650. The initial parameter vector is

$$\underline{p} = [-25.0 \ 0.3 \ -1.0 \ 0.0]^T$$

The optimally compensated system ε_{PGM} is 0.06369799. The final parameter values are

$$\underline{p} = [-3925 \ -22.78 \ -157.0 \ 0.002317]^T$$

A plot of the compensated nonlinear interference as well as the uncompensated one is given in Fig. 31. Notice that 20dB of interference reduction has been achieved with only a second-order compensator.

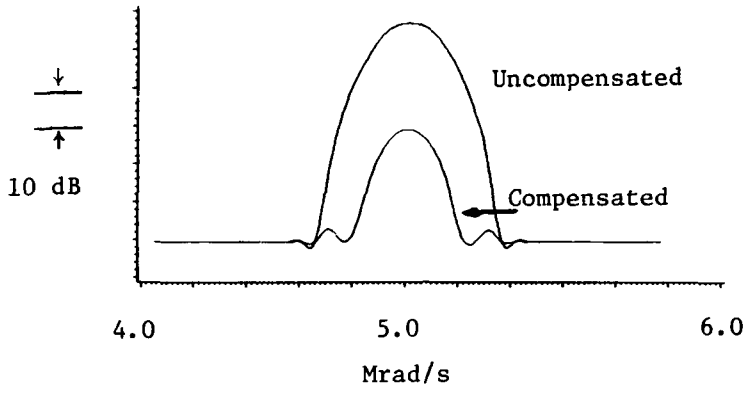


Fig. 31. Compensated versus the uncompensated interference.

Example 4

This example has a linear TF with a very narrow bandwidth. The linear characteristic of the channel is modeled by a sixth-degree Chebyshev with a passband of 4.9Mrad/s to 5.1Mrad/s.

$$H_1(s) = \frac{0.005708s^3}{s^6 + 0.250338s^5 + 75.03610s^4 + 12.51843s^3 + 1875.272s^2 + 156.3556s + 15609}$$

The initial parameter vector is

$$\underline{p} = [-38.07\ 0.2547\ 1.522\ 0]^T$$

The uncompensated system ε_{PGM} is 68.5240325. The final parameter values are

$$\underline{p} = [11390\ 37.12\ 455.8\ 2.2]^T$$

The system MSE is 21.5367901 representing a narrow improvement over the

uncompensated case. The compensated nonlinear interference is compared with the uncompensated one in Fig. 32

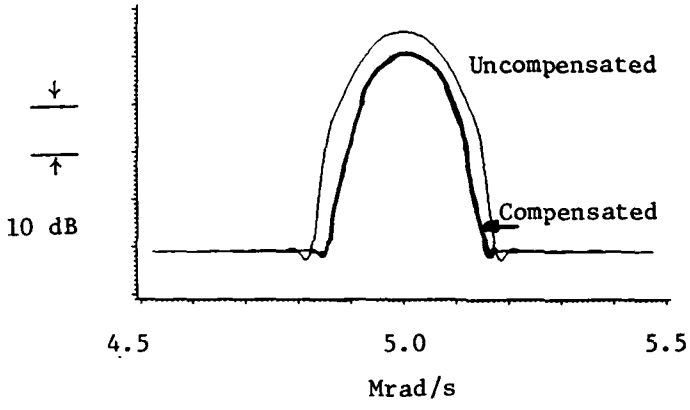


Fig. 32. Compensated versus the uncompensated interference.

CHAPTER V

CONCLUSIONS

The performance of a communication system can suffer degradation due to the intermodulation (IM) interference created by the inherent nonlinearities. It has been shown here that such IM interference can be effectively reduced for the frequency band(s) of interference by a stable post-compensator appended to the System Under Compensation (SUC). A complete methodology for designing such a compensator has been developed. It is based upon a mean-square IM criterion and was successfully implemented in a FORTRAN program VCOMP3. This program uses a highly powerful software package for optimization of mean-square nonlinear functions, thereby ensuring convergence to the global minimum in almost all practical cases.

The key strengths of the new approach are (a) broadband compensation (over the band of interest) in contrast with the single frequency compensation attempted in the past, and (b) guaranteed stability of the compensator. The latter was achieved in a highly innovative way. The stable region of the space of the parameters (of block transfer function of the compensator) was mapped onto a new parameter space. In fact the stable region was the interior of the s-plane to the left of the constant damping ratio lines, with a damping ratio ζ specified by the designer. Thus for the reparameterized problem the program VCOMP3 is unconstrained to choose any point in the new

space. The final design is not only stable, the block transfer functions have a damping ratio greater than ζ.

The compensator structures proposed, and used, have blocks with pole-zero transfer functions, hence only low order blocks are required. Indeed as shown in the studies presented, only first or second-order blocks were required. The concept of group delay for the interference component of the SUC was defined, which was followed by an improved compensator design. The key advantage realized is that the dimensionality of the parameter vector was small, typically 3 to 9, depending on the particular structure. Several case examples were presented which demonstrate that reductions of 15 to 50 dB in the IM interference can be achieved.

For completeness of this volume, a fairly comprehensive discussion of Volterra systems was included. This research clearly shows that advanced theory can be utilized to yield a practical tool for improving the performance of expensive communication systems.

Future work can entail a) simultaneous IM interference in several bands, and b) simultaneous IM interference reduction for multiple receivers.

REFERENCES

[1] M. Schetzen, "The Volterra & Wiener Theories of Nonlinear Systems", New York: John Wiley & Sons, 1980.

[2] V. K. Jain, A. M. Bush, and D. J. Kenneally, "Volterra Transfer Functions from Pulse Tests for Mildly Nonlinear Channels", Technical Report, RADC-TR-83-157, Rome Air Development Center, July 1983.

[3] V. K. Jain, S. J. Garrett, and A. R. Gondeck, "Analytical Methods for Characterization of Nonlinear Devices and Networks", Technical Report, RADC-TR-85-243, Vol I, Rome Air Development Center, December 1985.

[4] J. E. Dennis, Jr., D. M. Gay, and R. E. Welsch, "An Adaptive Nonlinear Least-Squares Algorithm", Technical Report, TR-20, Massachusetts Institute of Technology, October 1980.

[5] S. Narayanan, "Application of Volterra Series to Intermodulation Distortion Analysis of a Transistor Feedback Amplifier", IEEE Transactions on Circuit Theory, Vol 17(4), pp. 518-527, November 1970.

[6] R. P. Hecken, and R. C. Heidt, "Predistortion Linearization of the AR 6A Transmitter", Proc. International Conference on Comm., Vol 2, pp. 33.1.1-33.1.6, 1980.

[7] J. E. Dennis, and J. J. More, "Quasi-Newton Methods, Motivation and Theory", SIAM Review, pp. 46-89, January 1977.

[8] J. E. Dennis, "Nonlinear Least Squares and Equations", London: Academic Press, 1977.

[9] C. G. Broyden, "Quasi-Newton Methods and their Application to Function Minimization", in Math Comp., Vol. 21, pp. 368-381, 1967.

[10] M. S. Bazaraa and C. M. Shetty, "Nonlinear Programming Theory and Algorithms", New York: John Wiley and Sons, 1979.

[11] D. G. Luenberger, "Introduction to Linear and Nonlinear Programming", Reading, Mass: Addison-Wesley, 1973.

[12] V. K. Jain, and T. E. McClellan, "Stable Compensation of Nonlinear Communications Systems (Using Volterra Systems Characterization)", Technical Report, RADC-TR-85-243, Vol II, Rome Air Development Center, December 1985.

*[13] V. K. Jain and T. E. McClellan, "A Computer Program for the Design of Compensators for Nonlinear Communications Systems", RADC-TR-85-243, Vol III, Technical Report, Rome Air Development Center, December 1985.

* Although this report references the above limited document, no limited information has been extracted. Distribution on this document is limited to US Government agencies and their contractors; critical technology; Dec 85. Other requests for this document shall be referred to RADC (RBCT), Griffiss AFB, NY 13441-5700.

APPENDIX

APPENDIX

SIMPLE LINEAR COMPENSATION EXAMPLE

In this section we offer a tutorial example in altering linear system response. This material is essential to the compensating of H_1 and it provides us with a fundamental compensator structure to reduce the effects of the higher order terms.

These examples deal with the most rudimentary adjustment of linear distortion in a communication channel. In both examples we take the general form of the compensator to be

$$G_1(s) = \frac{b_1 s + b_0}{s^2 + a_1 s + a_0}$$

Example 1

Consider a communication system characterized by the linear transfer function

$$H_1(s) = \frac{4(10^{-5})s^2}{s^4 + (0.283)s^3 + (2.02)s^2 + (0.28)s + 9.8}$$

where all frequencies are scaled by a factor of 10^6. The frequency band of interest is

$$\langle R \rangle = \langle (\tfrac{1}{2})0.9(10^6)\pi \ (\tfrac{1}{2})1.1(10^6)\pi \rangle$$

$$f_0 = (\tfrac{1}{2})10^6\pi \text{ Hz}$$

The system frequency characteristics are given in Table 3. As a measure of the effect of the compensator on the phase characteristic, a linear approximation to the phase of each example will be given.

Linear Approximation to the slope

$$m\cdot\omega + c$$

$$m = -0.889\cdot10^{-3}$$

$$c = -4.289$$

Maximum phase deviation from a linear approximation

$$3.367^{o}$$

TABLE 3

THE UNCOMPENSATED SYSTEM FREQUENCY CHARACTERISTICS

$\omega \times 10^6$	Magnitude	Phase(degrees)
0.90	0.177	89.87
0.95	0.244	39.30
1.00	0.249	-4.28
1.05	0.239	-46.81
1.10	0.176	-89.91

It is desired to provide a second-order post compensator to flatten the magnitude characteristic over this band. The compensator structure chosen because of its wide bandwidth coverage is

$$G_1(s) = K\frac{s^2 + b_1 s + \omega_0^2}{(s + (0.1)\omega_0)(s + (10)\omega_0)}$$

K is chosen to provide unit gain at the center frequency ω_0. Thus b_1 is the only free design parameter. By experimentation a satisfactory value was found to be $b_1 = 200.0$, so that K = 50.64. The corresponding magnitude and phase of the compensator and the corrected system are given in Table 4. and Table 5.

TABLE 4

THE COMPENSATED SYSTEM FREQUENCY CHARACTERISTICS

The Compensator Response

$\omega \times 10^6$	Magnitude	Phase(degrees)
0.90	5.686	-43.92
0.95	4.464	-25.46
1.00	4.011	0.00
1.05	4.423	24.37
1.10	5.420	41.19

The final system response is shown in Table 5, overall response is a flat pass-band.

Linear Approximation to the slope

$m \cdot \omega + c$

$m = -0.433 \cdot 10^{-3}$

$c = -4.289$

Maximum phase deviation from a linear approximation

1.136^{o}

TABLE 5

THE COMPENSATED SYSTEM FREQUENCY CHARACTERISTICS

The Final System Response

$\omega \times 10^6$	Magnitude	Phase(degrees)
0.90	1.008	45.92
0.95	1.092	13.83
1.00	1.000	-4.28
1.05	1.058	-22.43
1.10	0.994	-48.72

The compensator parameter values which provided the desired results are listed below:

$a_0 = 1.0$

$a_1 = 10.1$

$b_0 = 1930.0$

$b_1 = 405.0$

$b_2 = 193$

$K = 50.64$

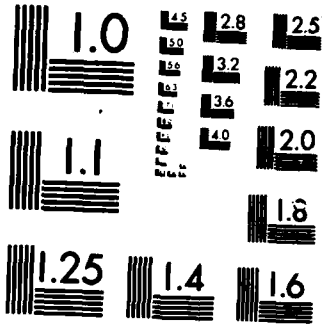
Several observations can be made concerning this example; first, the compensated magnitude characteristic is quite flat. Indeed, it is found that the maximum deviation from 0dB is only 0.2dB. Second, note that the maximum deviation of the uncompensated phase from a linear approximation given by

$$\Phi_H(\omega) = -0.89066(10^{-3})\omega - 4.28 \quad \text{(in degrees)}$$

is 3.43°. For the compensated system, the maximum phase deviation from

AD-A165 838 STABLE COMPENSATION OF NONLINEAR COMMUNICATIONS SYSTEMS (USING VOLTERRA S..(U) UNIVERSITY OF SOUTH FLORIDA TAMPA DEPT OF ELECTRICAL ENGINEER.. V K JAIN ET AL. 2/2

UNCLASSIFIED DEC 85 RADC-TR-85-243-VOL-2 F/G 17/2 NL



the linear approximation

$$\Phi_{HG}(\omega) = (0.015326)\omega$$

is 0.878^{o}.

Example 2

We consider the same communication system as before however, here the objective is to seek a compensator to reduce the 3dB band-width from 200 Krad/s to 100 Krad/s. A compensator structure is chosen that will narrow the pass-band it is

$$G_1(s) = K \frac{(s + (0.1)\omega_0)}{s^2 + a_1 s + \omega_0^2}$$

As before K is chosen to provide unit gain at the center frequency ω_0. The magnitude and phase of the compensator is given in Table 6. The magnitude and phase of the compensated system is given in Table 7.

TABLE 6

THE COMPENSATED SYSTEM FREQUENCY CHARACTERISTICS

The Compensator Response

$\omega \times 10^6$	Magnitude	Phase(degrees)
0.90	1.813	61.92
0.95	2.894	43.22
1.00	4.010	0.00
1.05	2.950	-42.09
1.10	1.940	-59.87

The system overall is a 3dB narrower pass-band.

Linear Approximation to the slope

$m \cdot \omega + c$

$m = -1.600 \cdot 10^{-3}$

$c = -4.289$

Maximum phase deviation from a linear approximation

14.49^{o}

TABLE 7

THE COMPENSATED SYSTEM FREQUENCY CHARACTERISTICS

The Final System Response

$\omega \times 10^{6}$	Magnitude	Phase(degrees)
0.90	0.321	151.80
0.95	0.708	82.53
1.00	1.000	-4.28
1.05	0.705	-88.90
1.10	0.343	-152.01

The compensator parameter values which provided the desired results are listed below:

$a_0 = 1.0$

$a_1 = 0.107$

$b_0 = 0.0421$

$b_1 = 0.425$

$b_2 = 0.421$

$K = 50.31$

Several observations can be made concerning this second example; first, the compensated magnitude characteristic reaches the -3dB points in half

the bandwidth required in the original system. Second, note that the maximum deviation of the uncompensated phase from a linear approximation given by

$$\Phi_H(\omega) = -1.600(10^{-3})\omega - 4.28 \quad \text{(in degrees)}$$

is 14.4°.

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.

END
FILMED
4-86
DTIC